Michael Rohlfing

Quasiteilchen-Bandstrukturen von Halbleitern und Halbleiter-Oberflächen

– 1996 –

Theoretische Physik

Quasiteilchen-Bandstrukturen von Halbleitern und Halbleiter-Oberflächen

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Physik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von
Michael Rohlfing
aus Minden
– 1996 –

Dekan:	Prof. Dr. F.-K. Holtmeier
Erster Gutachter:	Prof. Dr. J. Pollmann
Zweiter Gutachter:	Prof. Dr. C. Falter
Tag der mündlichen Prüfung(en):	.....
Tag der Promotion:	.....

Inhaltsverzeichnis

1 Einleitung **1**

I Einteilchen-Anregungen 9

2 Die GW-Näherung **13**
2.1 Green-Funktion und Selbstenergie-Operator 13
2.2 Quasiteilchen-Korrekturen 18

3 Darstellung von Zweipunkt-Funktionen **25**
3.1 Ebene Wellen und Gauß-Orbitale 25
3.2 Polarisierbarkeit 29
3.3 Dielektrische Matrix 30
3.4 Ergebnisse für Si 33

4 Das Plasmon-Pol-Modell **40**
4.1 Einfache Plasmon-Pol Modelle 41
4.2 Multi-Plasmon-Pol Modelle 46

5 Eigenschaften der Selbstenergie **53**

II Bandstrukturen von Halbleiteroberflächen 58

6 Das Slabkonzept **61**

7 Die Si(001)-(2×1) Oberfläche **70**
7.1 Das asymmetrische Dimer-Modell (ADM) 74
7.2 Die Quasiteilchen-Bandstruktur 78
7.3 Dielektrische Modellfunktionen 86
7.4 Das symmetrische Dimer-Modell (SDM) 91

8 Weitere reine Oberflächen **102**
8.1 Ge(001)-(2×1) 102

8.2 Vergleich der Ge(001)-(2×1) mit der Si(001)-(2×1) Oberfläche 110
8.3 β-SiC(001)-(2×1), C-terminiert 111

9 Adsorbat-bedeckte Oberflächen 118
9.1 H:Ge(001)-(2×1) 119
9.2 H:Si(111)-(1×1) 129
9.3 S:Ge(001)-(1×1) 134
9.4 As:Si(111)-(1×1) 140

10 Ionisierungsenergien 144

III Semicore- und Core-Elektronen in Halbleitern 147

11 Vorbemerkungen 148
11.1 Lokalisierte Zustände 148
11.2 Selbstwechselwirkungen 150

12 Technische Vorbemerkungen 153
12.1 Durchführung der LDA-Rechnungen 153
12.2 Das Pseudopotentialkonzept 156

13 Der Einfluß von Semicore-Elektronen ... 161
13.1 CdS als Prototyp 163
13.2 Analyse der Selbstenergien in CdS 176
13.3 Weitere Verbindungshalbleiter 183
13.4 Elementhalbleiter 191

14 Semicore- und Core-Niveaus 202
14.1 LDA und Standard-GWA 203
14.2 Erweiterungen der GWA 205

15 Core-Niveau Verschiebungen an Oberflächen 217
15.1 Die As:Si(111)-(1×1) Oberfläche 219
15.2 Die H:Si(111)-(1×1) Oberfläche 224
15.3 Die reine Si(001)-(2×1) Oberfläche 229

16 Zusammenfassung 233

Anhang **236**

A Darstellung von Zweipunkt-Funktion **236**

B Die Ansatzfunktionen **240**

C Rekursive Bestimmung von Matrixelementen **244**

D Dielektrische Matrix und Coulomb-Wechselwirkung **246**

E Coulomb-Divergenz in der Brillouinzonen-Integration **249**

F Parameterbestimmung in Funktionen mit Polstruktur **257**

Literaturverzeichnis **260**

Wesentliche Teile dieser Arbeit wurden bereits in den folgenden Referenzen publiziert.

- M. Rohlfing, P. Krüger und J. Pollmann,
  Quasiparticle band-structure calculations for C, Si, Ge, GaAs, and SiC using Gaussian-orbital basis sets,
  Phys. Rev. B **48**, 17 791 (1993).

- M. Rohlfing, P. Krüger und J. Pollmann,
  Efficient scheme for GW quasiparticle band-structure calculations with applications to bulk Si and to the Si(001)-(2×1) surface,
  Phys. Rev. B **52**, 1905 (1995).

- M. Rohlfing, P. Krüger und J. Pollmann,
  Quasiparticle Band Structure of CdS,
  Phys. Rev. Lett. **75**, 3489 (1995).

- M. Rohlfing, P. Krüger und J. Pollmann,
  Metallic nature of the symmetric dimer model of Si(001)-(2×1),
  Phys. Rev. B **52**, 13 753 (1995).

- M. Sabisch, P. Krüger, A. Mazur, M. Rohlfing und J. Pollmann,
  First-principles calculations of β-SiC(001) surfaces,
  Phys. Rev. B **53**, 13 121 (1996).

- J. Pollmann, P. Krüger, M. Rohlfing, M. Sabisch und D. Vogel,
  Ab-initio calculations of structural and electronic properties of prototype surfaces of group IV, III-V and II-VI semiconductors,
  Appl. Surf. Sci. **104/105**, 1 (1996).

- M. Rohlfing, P. Krüger und J. Pollmann,
  Quasiparticle band structures of clean, hydrogen- and sulfur-terminated Ge(001) surfaces,
  Phys. Rev. B **54**, 15. November (1996).

- M. Rohlfing, P. Krüger und J. Pollmann,
  Quasiparticle calculations for bulk semiconductors and their surfaces,
  in *Proceedings of the 23rd International Conference on the Physics of Semiconductors*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore 1996), p. 297.

Kapitel 1

Einleitung

Gegenstand der vorliegenden Arbeit ist die Untersuchung der elektronischen Eigenschaften kristalliner Festkörper. Dabei stehen die Energien von Einteilchen-Anregungen im Mittelpunkt. Darunter verstehen wir diejenigen Energien, die frei werden oder aufgewendet werden müssen, wenn einzelne Elektronen in das neutrale System hineingebracht oder daraus entfernt werden. Die Gesamtheit dieser Energien definiert die elektronische Bandstruktur des Kristalls. Einteilchen-Anregungsprozesse stellen eine einfache Möglichkeit dar, durch Beobachtung von außen Information über die Physik des Vielelektronensystems im Festkörper zu gewinnen. Typische Beispiele für die Realisierung dieser Anregungen sind Experimente der Photoemission und der inversen Photoemission. Mit diesen Techniken lassen sich Bandstrukturen mit hoher Präzision messen. Wir wollen in dieser Arbeit einen theoretischen Zugang vorstellen, der es ermöglicht, die Bandstrukturen mit der gleichen hohen Präzision zu berechnen. Der Zugang wird als ab-initio Methode formuliert.

Im Gegensatz zu empirischen Zugängen verzichten ab-initio Methoden auf eine vereinfachte Modellbildung und benötigen keine Vorgabe experimenteller, systemspezifischer Daten. Statt dessen werden die zu untersuchenden Größen, ausgehend von grundlegenden physikalischen Gesetzen und Regeln (first principles), berechnet. Dadurch decken ab-initio Methoden einen wesentlich weiteren Gültigkeitsbereich ab als empirische Modelle. Sie ermöglichen auch Prognosen für Systeme und Fragestellungen, die experimentell noch nicht untersucht werden konnten. Allerdings erzwingt bei einer solchen Vorgehensweise die Komplexität der Vielteilchenphysik eine Beschränkung auf diejenigen Teilaspekte der zugrunde liegenden Prinzipien, die für das jeweilige Problem relevant sind. Im Zusammenhang mit unserer Fragestellung nach der elektronischen Bandstruktur sind zwei spezielle ab-initio Methoden

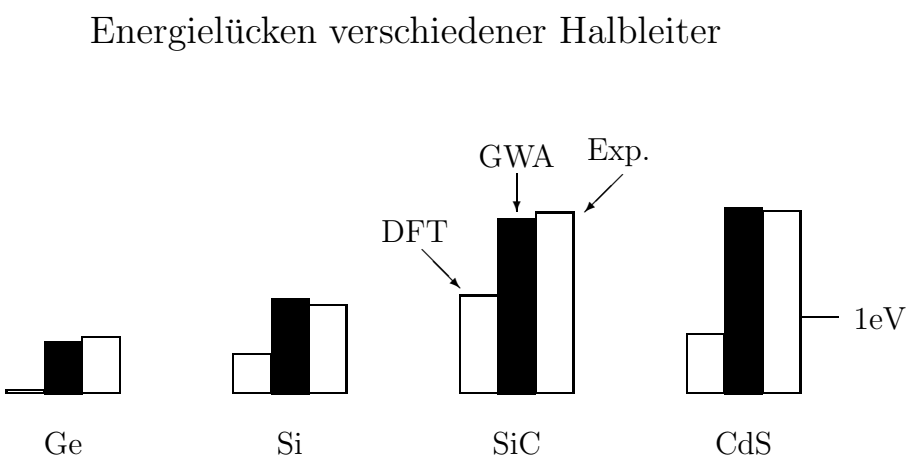


Abbildung 1.1: *Fundamentale Energielücken verschiedenen Volumenhalbleiter. Wir vergleichen DFT-LDA Bandstruktur-Ergebnisse (jeweils links) und GW Quasiteilchen-Gapenergien (Mitte) mit experimentell bestimmten Gapenergien.*

von besonderer Bedeutung: die Dichtefunktionaltheorie und die GW-Näherung.

Die Entwicklung der Dichtefunktionaltheorie (DFT) durch Hohenberg und Kohn [Hoh64] sowie durch Kohn und Sham [Koh65] stellt einen Meilenstein in der ab-initio Beschreibung von Festkörpern dar. In Form der lokalen Dichte-Approximation (LDA) ermöglicht die DFT eine quantitativ zuverlässige Bestimmung aller strukturellen Eigenschaften von Materie auf der mikroskopischen Längenskala der chemischen Bindung. Darüber hinaus ergeben sich im Rahmen der DFT energieartige Parameter, die häufig als elektronische Bandstruktur interpretiert werden. Diese DFT-Bandstrukturen weisen allerdings folgende systematische Abweichungen von gemessenen Bandstrukturdaten auf:

- Die fundamentale Energielücke von Halbleitern und Isolatoren wird unterschätzt.
- Die Bandbreite einfacher Metalle wird überschätzt.
- Die Bindungsenergien von Core- und Semicore-Elektronen werden unterschätzt.

Der erste Punkt, das *Gap-Problem der DFT*, stellt die bekannteste dieser systematischen Schwächen dar. Wir vergleichen in Abb. 1.1 für einige ausgewählte Halbleiter die berechnete DFT-LDA-Gapenergie (jeweils links) mit der experimentellen Gapenergie (jeweils rechts). Bei halbleitenden Systemen ist gerade die Größe der

Energielücke von entscheidender Bedeutung für eine Vielzahl von Anwendungen. Das Gap-Problem bildet eine entscheidende Schwachstelle bei der Interpretation der DFT-Energien als elektronische Bandstruktur.

Die systematischen Fehler der DFT-Bandstrukturen lassen sich als Folgen der konzeptionellen Gestaltung der DFT identifizieren. Der Begriff der Bandstruktur ist untrennbar mit der Betrachtung teilchenartiger Anregungen des Elektronensystems, also mit dem Hinzufügen oder Entfernen von Elektronen, verbunden. Solche Anregungen werden aber im Rahmen von DFT-Bandstrukturberechnungen gar nicht betrachtet. Die DFT beschreibt prinzipiell nur den Grundzustand des elektronischen Systems. Sie ist daher zur Berechnung von Bandstrukturen streng genommen gar nicht geeignet. [Wir werden später erläutern, warum die DFT-Energien dennoch in vielen Fällen recht gute Approximationen für die Bandstruktur darstellen.] Lediglich die fundamentale Energielücke könnte im Prinzip durch Vergleich der DFT-Grundzustandsenergien des Systems mit N, N–1 und N+1 Elektronen bestimmt werden.

Zur quantitativ richtigen Bestimmung von Bandstrukturen ist es notwendig, zu einem Konzept überzugehen, das explizit die teilchenartigen Anregungen behandelt. Die Vielteilchentheorie stellt solch ein Konzept zur Verfügung. Die zentrale Größe ist dabei die Green-Funktion der Elektronen. Sie beschreibt die Propagation eines zusätzlichen Elektrons oder Loches im fraglichen System und die damit verbundenen Einteilchen-Anregungs-Energien, die die *Quasiteilchen-Bandstruktur* definieren. [Die Bezeichnung als Quasiteilchen-Bandstruktur deutet an, daß die teilchenartige Propagation nicht mehr einem einzelnen Elektron zugeordnet werden kann. Statt dessen verteilt sich der teilchenartige Charakter der Anregung auf das Gesamtsystem und bildet ein Quasiteilchen.] Die Vielteilchentheorie gibt eine Bewegungsgleichung für die Green-Funktion an. Der darin enthaltene Selbstenergie-Operator, der die Austausch- und Korrelationseffekte im Elektronensystem beschreibt, besitzt eine äußerst komplizierte Struktur. Dadurch gestalten sich die Lösung der Bewegungsgleichung und die Berechnung von Bandstrukturen als sehr schwierig. Konkrete Resultate konnten mit diesem Formalismus erst nach Einführung einer geeigneten Näherung für den Selbstenergie-Operator erzielt werden. Die *GW-Näherung (GWA)*, die von Hedin und Lundqvist [Hed65, Hed69] vorgestellt wurde, führte hier zum entscheidenden Durchbruch. Allerdings beschränkte sich die Anwendbarkeit der GWA wegen des enormen numerischen Aufwandes zunächst auf homogene Systeme [Hed69, Pic84]. Erst die zunehmende Leistungsfähigkeit elektronischer Rechner eröffnete vor

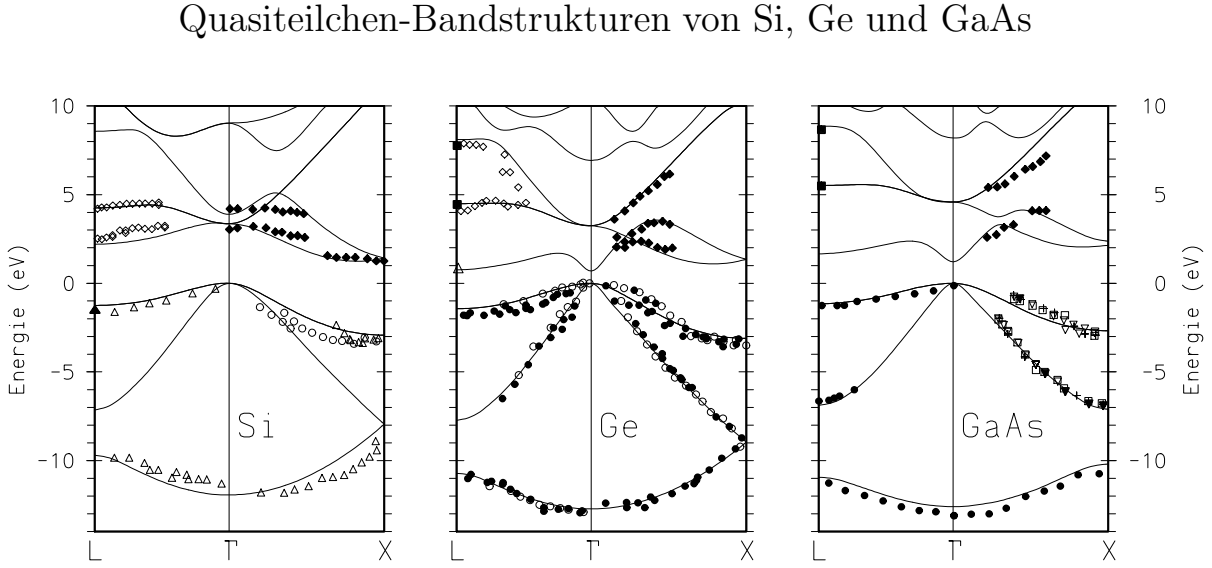


Abbildung 1.2: *Berechnete GW Quasiteilchen-Bandstrukturen der Volumenkristalle Si, Ge und GaAs. Die einzelnen Symbole stellen gemessene Bandstrukturenergien aus ARPES und ARIPES Experimenten dar (siehe [Ort93]).*

etwa zehn Jahren die Möglichkeit, die GWA auch auf inhomogene Systeme anzuwenden und Bandstrukturen realer Kristalle zu bestimmen [Hyb86, God88]. Dabei wird die GWA in der Regel als Korrektur zur DFT formuliert. Diese Vorgehensweise erweist sich als außerordentlich erfolgreich. In Abb. 1.2 stellen wir berechnete GW Quasiteilchen-Bandstrukturen von Si, Ge und GaAs [Roh93a] einer Vielzahl gemessener Bandstrukturdaten [Ort93] gegenüber. Die fundamentalen Gapenergien der GW Quasiteilchen-Bandstrukturen werden in Abb. 1.1 mit den zugehörigen DFT-Gapenergien und den experimentellen Werten verglichen. Die hervorragende Übereinstimmung der berechneten Daten mit dem Experiment demonstriert in beeindruckender Weise die Aussagekraft der GWA und ihr Potential für die Bestimmung elektronischer Anregungsenergien. Das Gap-Problem der DFT wird durch diesen Zugang vollständig überwunden.

In der Folgezeit wurde die GWA auf eine Vielzahl unterschiedlichster elektronischer Systeme angewendet. Das ursprüngliche, klassische Anwendungsfeld der GWA bilden die Volumenhalbleiter der Gruppe IV (C, Si, Ge, SiC) und die III-V Verbindungen, die aus B, Al, Ga oder In sowie aus N, P, As oder Sb zusammengesetzt sind [Hyb86, God88, Lin88, Hot91, Zhu91b, Roh93a, Roh93b, Rub93, Pal94, Bac95, Bla95, Gry95, Wen95b]. Andere halbleitende Volumenkristalle lassen sich ebenfalls

behandeln. Hierzu zählen etwa die II-VI Halbleiter, die aus Zn oder Cd sowie aus O, S, Se oder Te zusammengesetzt sind [Zak94], die Alkalihalogenide, z. B. LiCl [Hyb86], sowie Übergangsmetalloxide, wie etwa NiO [Ary95]. Beispiele für die Untersuchung von Metallen sind Na, Li und Al [Nor87, Nor89], ferner Ag [Hor87] sowie das Übergangsmetall Ni [Ary92]. Auch das Energiespektrum einfacher Legierungen, etwa aus Si und C, kann mit der GWA bestimmt werden [Gry95].

Die Anwendung der GWA beschränkt sich aber nicht auf dreidimensional ausgedehnte Systeme. Mit der zunehmenden Leistungsfähigkeit elektronischer Rechner wurden auch Systeme mit gestörter Periodizität einer GW-Behandlung zugänglich. Im Rahmen der GWA wurden Bandstrukturen für verschiedene reine und adsorbatbedeckte Oberflächen von C, Si, Ge und GaAs bestimmt [Hyb87a, Hyb88b, Zhu89a, Nor91, Zhu91, Hri93, Kre93, Nor93, Bla94, Kre94, Kre95]. Hinzu kommen Untersuchungen an Si/Ge Supergittern [Hyb87b], an Grenzflächen zwischen Al und GaAs [Cha93] und an kristallinem C_{60} [Shi93a]. Ferner erlaubt die GWA auch die Berechnung des Spektrums von Atomen [Shi93b], von kleinen Molekülen und Clustern wie etwa Na_4 [Oni95], sowie von Defekten wie etwa O in Si [Liz95] oder F-Zentren in LiCl [Sur95].

Wir stellen dieses vielseitige Konzept der GW-Näherung zur Berechnung von Bandstrukturen in den Mittelpunkt unserer Arbeit. Unser Ziel ist die Untersuchung solcher Systeme, die sich mit den herkömmlichen Techniken zur Umsetzung der GW-Näherung nicht oder nur mit sehr hohem Aufwand behandeln lassen. Hierzu zählen insbesondere Systeme, in denen stark lokalisierte elektronische Zustände auftreten. In diesem Fall ist die geeignete Wahl der räumlichen Basisfunktionen für die numerische Realisierung der GWA von entscheidender Bedeutung. Basissätze aus ebenen Wellen, wie sie in den allermeisten Arbeiten verwendet werden, erweisen sich hier wegen der notwendigen großen Zahl an Ansatzfunktionen als unpraktikabel. Statt dessen stellen wir einen grundsätzlich neuen Zugang zur GWA und zur Berechnung von Quasiteilchen-Bandstrukturen vor. Er beruht auf der Verwendung von Basissätzen aus lokalisierten Gauß-Orbitalen, die sich gerade zur Beschreibung lokalisierter Zustände als äußerst vorteilhaft und effizient herausstellen. Die algebraischen Eigenschaften der Gauß-Orbitale unterscheiden sich wesentlich von denen der ebenen Wellen. Die Neuformulierung der GWA in dieser Basis stellt daher einen erheblichen Aufwand dar, der sich aber wegen der wesentlich geringeren Basisgröße in den numerischen Berechnungen als lohnenswert erweist.

Lokalisierte Gauß-Orbitale werden seit langem in der DFT erfolgreich zur Dar-

stellung der elektronischen Wellenfunktionen benutzt. Sie erlauben besonders bei Systemen, die lokalisierte elektronische Zustände aufweisen, eine effiziente Durchführung der DFT. Wir haben in [Roh93a, Roh93b] einen Zugang vorgestellt, der auf der Basis solcher DFT-Rechnungen mit Gauß-Orbitalen die Bestimmung der Quasiteilchen-Bandstruktur im Rahmen der GWA erlaubt. Dabei wurden allerdings die Zweipunkt-Funktionen, die in der GWA auftreten, nach wie vor mit ebenen Wellen dargestellt. In der vorliegenden Arbeit drücken wir nun auch diese Zweipunkt-Funktionen durch lokalisierte Gauß-Orbitale aus. Diese Darstellungstechnik wird hier detailliert vorgestellt.

Durch die Verwendung von Gauß-Orbitalen reduziert sich bei vielen Systemen der numerische Aufwand der GWA gegenüber einer Ebenen-Wellen Rechnung erheblich. Dies eröffnet uns den Zugang zur Behandlung zweier zunächst sehr unterschiedlicher physikalischer Problemklassen, die sich beide mit Hilfe der Gauß-Orbitale effizient behandeln lassen.

- In Teil II unserer Arbeit untersuchen wir reine und adsorbatbedeckte Halbleiteroberflächen. Die geänderte chemische Umgebung der Oberflächenatome und der Übergang vom Kristall zum Vakuum führen zum Auftreten elektronischer Zustände, die an der Oberfläche lokalisiert sind. An reinen Halbleiteroberflächen bilden die nichtabgesättigten, freien Bindungen (Dangling Bonds) das markanteste Merkmal der elektronischen Struktur. Durch die Absättigung mit Adsorbatatomen werden sie in chemische Bindungen umgewandelt. Wir werden die lokalisierten elektronischen Oberflächenzustände in den Mittelpunkt der Betrachtungen rücken und ihre Quasiteilchen-Energien bestimmen.

Die Oberfläche eines Kristalls zerstört die Gitterperiodizität in senkrechter Richtung. Daher sind GW Rechnungen für Oberflächen um ein Vielfaches aufwendiger als für Volumenkristalle. Die Verwendung von Gauß-Orbitalen ermöglicht bei vielen Oberflächen eine erhebliche Reduktion dieses großen numerischen Aufwandes.

- In Teil III der Arbeit wenden wir uns den Eigenschaften von lokalisierten Core- und Semicore-Elektronen in Kristallen zu. Diese Zustände scheinen auf den ersten Blick auf die Kristalleigenschaften (chemische Bindungen, Bandstruktur, optische Eigenschaften) keinen Einfluß zu haben. Dieser Eindruck täuscht jedoch. Wir werden zu dem überraschenden Ergebnis gelangen, daß die Semicore-Elektronen in vielen Kristallen, insbesondere in den II-VI Halbleitern, einen erheblichen Einfluß auf die Quasiteilchen-Bandstruktur ausüben

und zu einer konsistenten Beschreibung keinesfalls vernachlässigt werden dürfen. Wir untersuchen diese Einflüsse bei verschiedenen Verbindungs- und Elementhalbleitern. Daraus lassen sich systematische Trends bezüglich der Bedeutung der Semicore-Zustände ableiten, die sich aus der Stellung der beteiligten Elemente im Periodensystem ergeben.

Neben ihrem Einfluß auf die Kristalleigenschaften betrachten wir auch die Energien der lokalisierten Zustände selbst. Wir hatten erwähnt, daß ihre Bindungsenergie in DFT-Bandstrukturen systematisch unterschätzt wird.

Die GWA wird es uns ermöglichen, diese Bindungsenergie wesentlich genauer zu bestimmen. Ferner untersuchen wir die Verschiebung der Core-Niveaus an Halbleiter-Oberflächen durch die dort geänderte chemische Umgebung der Substratatome. Diese Untersuchung stellt eine Verknüpfung zwischen der Oberflächenphysik aus Teil II und der Energetik der Core-Niveaus dar.

Bei der Untersuchung der stark lokalisierten Semicore- und Core-Zustände treten die Vorteile der Gauß-Orbitale gegenüber den ebenen Wellen als Basisfunktionen offenkundig zutage. Bei Verwendung der Gauß-Orbitale lassen sich die Wellenfunktionen der Zustände sehr einfach durch entsprechend steile Ansatzfunktionen erfassen. Wenn sie mit gleicher Genauigkeit mit ebenen Wellen beschrieben werden sollten, so müßte deren Abschneide-Energie weit größer gewählt werden als etwa bei der Behandlung gewöhnlicher Halbleiter. Die Beschreibung von Core- und Semicore-Zuständen mit Hilfe ebener Wellen erfordert daher um Größenordnungen umfangreichere Basissätze und längere Rechenzeiten. Ohne eine lokalisierte Basis wäre die Untersuchung dieser Zustände zur Zeit überhaupt nicht denkbar.

Die Arbeit ist wie folgt gegliedert. Im ersten Teil formulieren wir die theoretischen Grundlagen unserer Rechnungen. Wir werden zunächst das Konzept der GW-Näherung kurz vorstellen (Kap. 2). In Kap. 3 erläutern wir, wie die Zweipunkt-Funktionen der GWA durch lokalisierte Gauß-Orbitale dargestellt werden. Zur Kontrolle der Leistungsfähigkeit und Genauigkeit der Methode berechnen wir die Quasiteilchen-Bandstruktur von Si und vergleichen sie mit einer gewöhnlichen Ebenen-Wellen Rechnung. In Kap. 4 stellen wir die Plasmon-Pol Modelle vor, mit deren Hilfe sich die Frequenzintegrale des Selbstenergie-Operators behandeln lassen. Eine Diskussion der Eigenschaften von Selbstenergien elektronischer Zustände beendet diese Einführung in die GWA (Kap. 5).

Im zweiten Teil der Arbeit wenden wir das erarbeitete Konzept auf verschie-

dene Halbleiter-Oberflächen an und bestimmen die Quasiteilchen-Bandstrukturen von lokalisierten Oberflächenzuständen. Wir erläutern zunächst das Slab-Konzept (Kap. 6), das eine Ersatzgeometrie der realen Oberfläche darstellt und ihre konkrete Behandlung ermöglicht. Als erste Oberfläche untersuchen wir die Si(001) Oberfläche (Kap. 7) als ein typisches Beispiel für eine reine Halbleiteroberfläche. In Analogie dazu behandeln wir in Kap. 8 die reinen Ge(001) und SiC(001) Oberflächen. In Kap. 9 werden die freien Bindungen der reinen Oberflächen durch Fremdatome abgesättigt. Wir untersuchen die Adsorption von H, S und As auf Ge(001) und Si(111) Oberflächen. Ferner diskutieren wir in Kap. 10 die Ionisierungsenergien von Halbleiter-Oberflächen.

Im dritten Teil unserer Arbeit wenden wir uns den Eigenschaften stark lokalisierter Core- und Semicore-Zustände und ihrer Bedeutung für die Festkörpereigenschaften zu. Nach einigen technischen Anmerkungen zur Durchführung der zugrunde liegenden DFT-LDA im Falle stark lokalisierter Zustände und zur Konstruktion von Pseudopotentialen (Kap. 12) diskutieren wir in Kap. 13 die Bedeutung von Semicore-Elektronen für die Berechnung von Quasiteilchen-Bandstrukturen. Dabei stellen wir den II-VI Halbleiter CdS als Prototyp in den Vordergrund. In gleicher Weise untersuchen wir andere Verbindungs- und Elementhalbleiter. Abschließend betrachten wir in Kap. 14 die Energie der Core- und Semicore-Zustände selber. Dazu ist es erforderlich, den GW-Formalismus gezielt zu erweitern. An Halbleiteroberflächen treten Verschiebungen der Core-Niveaus gegenüber den Volumen-Energien auf. Dieses Problem behandeln wir in Kap. 15. Unser Formalismus ermöglicht die direkte Bestimmung solcher Verschiebungen. Wir betrachten exemplarisch die Si $2p$ Niveaus an reinen und adsorbatbedeckten Si Oberflächen.

Wir beschließen diese Arbeit, die in ihren wesentlichen Teilen bereits publiziert ist, mit einer Zusammenfassung.

Teil I

Einteilchen-Anregungen

Dieser Abschnitt unserer Arbeit beschäftigt sich mit der quantenmechanischen Behandlung eines Vielelektronensystems. Als Folge der Korrelation zwischen den Elektronen unterscheidet sich diese Behandlung wesentlich von der Quantenmechanik eines einzelnen Teilchens. Als ununterscheidbare Spin-1/2-Teilchen sind die Elektronen durch eine antisymmetrische Vielteilchen-Wellenfunktion zu beschreiben. Daraus ergibt sich als unmittelbare Folge das Pauli-Verbot, demzufolge Elektronen gleichen Spins sich nicht am gleichen Ort aufhalten können: Die Wahrscheinlichkeitsdichte für das Antreffen zweier Elektronen am gleichen Ort $\mathbf{r}$ beträgt Null. Elektronen gleichen Spins können daher nicht unabhängig voneinander behandelt werden. Die einfachste Möglichkeit, der Antisymmetrie gerecht zu werden, besteht in der Konstruktion der Wellenfunktion in Form einer einzelnen Slater-Determinante. Dieser Ansatz führt zur Hartree-Fock-Theorie, die als wesentliche Ergänzung zur Einteilchen-Quantenmechanik einen Austauschenergie-Operator beinhaltet. Dabei bleibt allerdings noch unberücksichtigt, daß auch Elektronen verschiedenen Spins aufgrund ihrer gegenseitigen Coulomb-Abstoßung einer Korrelation unterliegen. Auch für sie wird die Wahrscheinlichkeitsdichte für den Aufenthalt zweier Elektronen am gleichen Ort gegenüber dem Produkt der Einteilchen-Aufenthaltswahrscheinlichkeiten reduziert sein. Der Begriff der Korrelation bedarf an dieser Stelle einer sprachlichen Präzisierung. Unter *Korrelation* verstehen wir in dieser Arbeit den Unterschied zwischen der Hartree-Fock-Theorie und der exakten Vielelektronenphysik. Sie wirkt also zusätzlich zum reinen Austausch-Prinzip (vgl. Tabelle 1.1).

Das Ziel dieser Arbeit ist die Beschreibung von Anregungen. Als *Einteilchen-Anregung* eines Festkörpers bezeichnen wir den Prozeß, die Elektronenzahl des N-Elektronensystems um ein Teilchen zu erhöhen (elektronische Anregung; $N \rightarrow N+1$) bzw. zu verringern (lochartige Anregung; $N \rightarrow N-1$). Aufgrund der Wechselwirkung zwischen den Elektronen kann eine solche elementare Anregung trotz ihres teilchenartigen Charakters nicht einem einzelnen Elektron zugeordnet werden; man bezeichnet sie deshalb als *Quasiteilchen*. Analog zu einem echten Teilchenzustand in einem wechselwirkungsfreien System ist ein Quasiteilchen durch Wellenfunktion (Quasiteilchen-Amplitude), Impuls und Energie charakterisiert. Diese Quasiteilchen-Energie wird bei einer Teilchenzahländerung des Systems frei bzw. muß aufgewendet werden; sie ist daher experimentell meßbar. Die Gesamtheit dieser Quasiteilchen-Energien definiert die elektronische Bandstruktur eines Festkörpers.

Unter der *Selbstenergie* eines elektronischen Einteilchen-Anregungszustandes ver-

	Hartree-Näherung	Hartree-Fock-Näherung	LDA	GW-Näherung	Vielteilchen-theorie
Einteilchen-Hamiltonian	×	×	×	×	×
Austausch-Prinzip		×	(×)	×	×
Korrelation				(×)	×

× = exakt behandelt (×) = approximativ berücksichtigt

Tabelle 1.1: *Berücksichtigung von Austausch- und Korrelationseffekten in verschieden Näherungs-Methoden der Vielteilchenphysik.*

steht man denjenigen Beitrag zu seiner Quasiteilchen-Energie, der von den Austausch-Korrelations-Effekten herrührt. [Die klassische Hartree-Energie zählt üblicherweise nicht dazu.] Den entsprechenden Term in der Bewegungsgleichung für den Anregungszustand bezeichnet man als *Selbstenergie-Operator*. Austausch und Korrelation zwischen den Elektronen führen wegen des abstoßenden Charakters der Coulomb-Wechselwirkung zu einer Absenkung der Energie gegenüber einer Beschreibung im Rahmen der Einteilchen-Quantenmechanik; dementsprechend ist die Selbstenergie grundsätzlich negativ. Die Selbstenergie der einzelnen Zustände liegt bei vielen Systemen in der Größenordnung von –5 bis –20 eV; für Core-Elektronen ergeben sich weit größere Werte im Bereich –100 eV. Bei räumlich ausgedehnten Systemen liegen sowohl der Austausch-Anteil als auch der Korrelations-Anteil der Selbstenergie in gleicher Größenordnung, so daß für eine quantitativ richtige Beschreibung keiner der beiden Anteile vernachlässigt werden kann.

Die *GW-Näherung* stellt eine äußerst erfolgreiche Methode zur approximativen Bestimmung des Selbstenergie-Operators dar. Sie liefert für eine Vielzahl verschiedenster Systeme eine zuverlässige Beschreibung der Austausch-Korrelations-Effekte und bietet somit die Möglichkeit, quantitativ richtige Quasiteilchen-Bandstrukturen, also Einelektronen-Anregungsspektren, zu berechnen. Die vollständige Darstellung eines Zuganges zur GW-Näherung würde den Rahmen dieses Abschnitts sprengen; wir verweisen hierzu auf die Literatur der Vielteilchentheorie, besonders auf die grundlegenden Arbeiten von Hedin und Lundqvist [Hed65, Hed69]. Nach einer kurzen Einführung wollen wir hier vielmehr die Anwendung dieser Methode auf inhomogene Halbleitersysteme diskutieren. Dabei sollen die Unterschiede zwischen unserem Zugang und den Arbeiten anderer Autoren [Hyb86, God88, Lin88, Hot91] im

Vordergrund stehen. Ein Schwerpunkt liegt in der Einführung lokalisierter Basisfunktionen von Gauß-Orbitalen anstelle ebener Wellen für die Darstellung der auftretenden Zweipunkt-Funktionen. Dieser Basissatz ermöglicht uns die Untersuchung stark lokalisierter Core- und Semicore-Elektronen sowie eine effiziente Behandlung von Oberflächensystemen.

Kapitel 2

Die GW-Näherung

Die Beschreibung der elektronischen Austausch-Korrelations-Effekte geschieht mit Hilfe eines Green-Funktions-Formalismus, wie er im Rahmen der Vielteilchen-Theorie aufgestellt wird [Hed65, Hed69]. Die GW-Näherung ist eine Näherung für den Selbstenergie-Operator innerhalb dieses Formalismus; sie ermöglicht die konkrete Umsetzung auf inhomogene Systeme.

2.1 Green-Funktion und Selbstenergie-Operator

Die zentrale Größe der Vielteilchen-Theorie ist die elektronische *Einteilchen-Green-Funktion*, die wir in ihrer kausalen, d.h. zeitgeordneten Form verwenden:

$$G(\mathbf{x}t, \mathbf{x}'t') = -i \left\langle N \left| T\left(\psi(\mathbf{x}, t)\psi^+(\mathbf{x}', t')\right)\right| N \right\rangle \quad . \tag{2.1}$$

Hierbei bezeichnet $|N\rangle$ den Vielteilchen-Grundzustand des Systems aus N Elektronen, $\psi(\mathbf{x}, t)$ den Einteilchen-Feldoperator (im Heisenberg-Bild), und T den Zeitordnungs-Operator. Die Variable $\mathbf{x}$=($\mathbf{r}$,σ) faßt Raum- und Spinkoordinaten zusammen.

Wir beschränken uns in dieser Arbeit grundsätzlich auf die Temperatur Null. Eine systematische Erweiterung der Vorgehensweise auf Temperaturen größer als Null ist mit Hilfe von Matsubara-Green-Funktionen prinzipiell möglich [Hed69].

Die Green-Funktion beschreibt die räumliche Propagation eines zusätzlichen Elektrons bzw. Loches im N-Elektronen-System im Laufe der Zeit. Für zeitlich homogene Systeme gelangt man mittels einer Fouriertransformation bezüglich der Zeitdifferenz $\tau = t - t'$ zu der energieabhängigen Green-Funktion

$$G(\mathbf{x}, \mathbf{x}', \omega) = \int G(\mathbf{x}t, \mathbf{x}'t')e^{i\omega\tau}d\tau \quad , \tag{2.2}$$

deren Pole den Anregungsenergien des Systems durch Entfernen oder Hinzufügen eines einzelnen Elektrons entsprechen. Die Green-Funktion steht also in direktem Zusammenhang mit den Meßergebnissen entsprechender Experimente, etwa der Photoemission, der inversen Photoemission und des elektronischen Tunneleffekts. Aber auch Experimente, die an elektronische Eigenschaften anknüpfen, ohne die Elektronenzahl zu verändern, sollten sich – wenn auch mit Einschränkungen – mit Hilfe der Einteilchen-Green-Funktion interpretieren lassen. Hierzu zählen optische Prozesse, wie etwa Ellipsometrie- und Reflektivitätsmessungen, und elektronische Streuexperimente, wie etwa niederenergetische Elektronenstreuung (LEED) [Hil95].

Wenden wir uns nun der Berechnung der Green-Funktion zu. Ausgangspunkt ist der Vielteilchen-Hamilton-Operator der Elektronen

$$H = \sum_{i=1}^{N} \left(-\nabla_i^2 + V_{\text{ext}}(\mathbf{r}_i) \right) + \sum_{i<j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad . \tag{2.3}$$

[Zur Vereinfachung der Schreibweise setzen wir $\hbar$=1, $\hbar^2/2m$=1 sowie e^2=2. Mit diesen Einheiten ergeben sich alle Längen als Vielfache des Bohr'schen Radius a_B=0.529177 Å und alle Energien als Vielfache der Rydberg-Einheit Ry=13.6058 eV. Ferner wird zwischen Energien und Frequenzen nicht unterschieden.] In dem Hamiltonian tritt das externe Potential V_{ext} auf, das in seiner räumlichen Struktur unmittelbar von der Position der Atome im System abhängt. Vor einer Behandlung des Elektronensystems muß also zunächst die atomare Geometrie feststehen. Hierzu gehört neben der Festlegung der Kristallstruktur die experimentelle oder theoretische Bestimmung der Gitterkonstanten. Falls Defekte, Grenzflächen oder Oberflächen untersucht werden, so muß zusätzlich die lokale Änderung der Geometrie, also das Rekonstruktions- und Relaxationsverhalten, festgestellt werden. Falls die atomare Struktur mikroskopisch berechnet werden soll, so geschieht das äußerst erfolgreich und verläßlich im Rahmen der DFT (LDA). Die quantenmechanischen Freiheitsgrade der Atomkerne bleiben im Hamiltonian (2.3) unberücksichtigt. Im Rahmen der Born-Oppenheimer-Näherung betrachten wir die Kerne als klassische Teilchen. Ihre Koordinaten treten lediglich in Form von Parametern in (2.3) auf. Ferner beschränken wir uns in dem Hamiltonian auf diejenigen Elektronen, die für die jeweils zu untersuchenden Eigenschaften des Systems relevant sind. Für die Bestimmung von herkömmlichen Bandstrukturen und von Oberflächenbandstrukturen sind das primär die Valenzelektronen. Soweit möglich, werden die energetisch tieferen Core-Elektronen im Rahmen eines Pseudopotential-Zugangs aus dem Formalismus

eliminiert; sie gehen dann in das externe Potential mit ein. In der vorliegenden Arbeit geschieht das durch Verwendung normerhaltender ab-initio Pseudopotentiale. Zum Problem, inwieweit auch Semicore-Elektronen eliminiert werden dürfen, verweisen wir auf Kapitel 13.

Ausgehend vom Hamiltonian (2.3), erhält man für die Green-Funktion G eine Bewegungsgleichung

$$\left\{ i\frac{\partial}{\partial t} + \nabla^2 - V_{\text{ext}}(\mathbf{r}) - V_{\text{Coul}}(\mathbf{r}) \right\} G(\mathbf{x}t, \mathbf{x}'t') + i \int \frac{e^2}{|\mathbf{r} - \mathbf{r}''|} G_2(\mathbf{x}''t, \mathbf{x}''t, \mathbf{x}t, \mathbf{x}'t') d\mathbf{x}'' = \delta(\mathbf{x}, \mathbf{x}')\delta(t, t') \quad , \tag{2.4}$$

die die Einteilchen-Green-Funktion G mit der Zweiteilchen-Green-Funktion G_2 verknüpft. Das Aufstellen weiterer Bewegungsgleichungen ergibt jeweils einen Zusammenhang zwischen der n-Teilchen-Green-Funktion mit einer höheren $(n+1)$-Teilchen-Green-Funktion. Das Problem wird dadurch so komplex, daß an eine tatsächliche Lösung nicht zu denken ist.

Statt dessen wird mit dem *Selbstenergie-Operator* eine zusätzliche Größe eingeführt, die über die Zweiteilchen-Green-Funktion definiert ist und in Gl. (2.4) eine Entkopplung bewirkt:

$$\left\{ i\frac{\partial}{\partial t} + \nabla^2 - V_{\text{ext}}(\mathbf{r}) - V_{\text{Coul}}(\mathbf{r}) \right\} G(\mathbf{x}t, \mathbf{x}'t') - \int \Sigma(\mathbf{x}t, \mathbf{x}''t'') G(\mathbf{x}''t'', \mathbf{x}'t') d\mathbf{x}'' dt'' = \delta(\mathbf{x}, \mathbf{x}')\delta(t, t') \quad . \tag{2.5}$$

Schon an dieser Stelle wird klar, daß der Selbstenergie-Operator Σ wegen der vereinfachten Abhängigkeit von $\mathbf{x}$ und t nicht die gesamte Information umfassen kann, die in der Zweiteilchen-Green-Funktion enthalten ist. Dies führt beispielsweise dazu, daß exzitonische Prozesse als typische Zweiteilchen-Korrelationseffekte zwischen Loch und Elektron mit diesem Selbstenergie-Operator nicht erfaßt werden können.

Der in Gl. (2.5) auftretende nichtlokale sowie zeit- bzw. energieabhängige Selbstenergie-Operator hängt selbstkonsistent mit der Green-Funktion zusammen. Seine exakte Bestimmung erweist sich als ebenso komplex und praktisch undurchführbar wie die Bestimung der höheren Green-Funktionen. Für eine konkrete Untersuchung ist man auf praktikable Näherungsausdrücke angewiesen.

Betrachten wir zunächst einen Ansatz, der als Vorstufe der GW-Näherung verstanden werden kann: $\Sigma = Gv$. Dieser Ausdruck für die Selbstenergie, der die

Green-Funktion G mit der reinen Coulomb-Wechselwirkung v verknüpft, stellt den ersten Term einer systematischen Entwicklung von Σ nach v dar. Er behandelt die Austausch-Effekte innerhalb des Elektronensystems richtig, vernachläsigt aber jegliche Korrelation. Diese Vorgehensweise ist identisch mit der Hartree-Fock-Theorie. Die praktische Durchführung ist technisch wesentlich einfacher als die der GW-Näherung. Es stellt sich jedoch heraus, daß die Hartree-Fock-Näherung für ausgedehnte Systeme Bandstruktur-Ergebnisse liefert, die erheblich von denen genauerer Näherungen und von experimentellen Daten abweichen. Die Ursache dafür liegt darin begründet, daß die Coulomb-Wechselwirkung v zwischen zwei Elektronen in einem ausgedehnten Elektronensystem erheblich *abgeschirmt* wird. Die resultierende *dynamisch abgeschirmte Wechselwirkung* $W(\mathbf{r}, \mathbf{r}', t-t')$ beschreibt die Wechselwirkung im Elektronensystem wesentlich besser als die reine Coulomb-Wechselwirkung $v(\mathbf{r}, \mathbf{r}')$. Eine entsprechende Reihenentwicklung des Selbstenergie-Operators nach dieser abgeschirmten Wechselwirkung W führt schon im ersten Term auf eine wesentlich bessere Beschreibung der Austausch-Korrelations-Effekte in Metallen und Halbleitern als die Hartree-Fock Theorie. Dieser erste Term lautet [Hed69]

$$\Sigma(\mathbf{x}, \mathbf{x}', t - t') = iG(\mathbf{x}, \mathbf{x}', t - t')W(\mathbf{x}, \mathbf{x}', (t - t') + 0^+) \quad . \tag{2.6}$$

Die durch (2.6) definierte *GW-Näherung* (GW Approximation, GWA) für den Selbstenergie-Operator wurde in den 60'er Jahren zunächst auf das Problem des homogenen Elektronengases angewendet [Hed65, Hed69], bevor in den 80'er Jahren die zunehmende Leistungsfähigkeit von elektronischen Rechnern die Anwendung auf reale, inhomogene Systeme ermöglichte [Hyb86, God88, Lin88].

Der erste Faktor im Ausdruck (2.6) ist durch die Einteilchen-Green-Funktion G gegeben. Der zweite Faktor, die abgeschirmte Wechselwirkung W, hängt im Rahmen der GW-Näherung in folgender Weise mit der Green-Funktion G zusammen. [Im folgenden bedeutet $(1)=(\mathbf{x}_1, t_1)$ etc.] Aus der Green-Funktion wird zunächst die Polarisierbarkeit

$$P(1, 2) = -iG(1, 2)G(2, 1) \tag{2.7}$$

berechnet. Eine anschließende Faltung mit der reinen Coulomb-Wechselwirkung v führt auf die dielektrische Funktion ϵ in der als Random-Phase-Approximation (RPA) bekannten Form:

$$\epsilon(1, 2) = \delta(1, 2) - \int v(1, 3)P(3, 2)d(3) \quad . \tag{2.8}$$

Durch Inversion erhält man die inverse dielektrische Funktion ϵ^{-1}, die mit ϵ wie folgt zusammenhängt:

$$\int \epsilon^{-1}(1,3)\epsilon(3,2)d(3) = \delta(1,2) \quad . \tag{2.9}$$

Eine erneute Faltung der inversen dielektrischen Funktion mit v ergibt die abgeschirmte Wechselwirkung:

$$W(1,2) = \int \epsilon^{-1}(1,3)v(3,2)d(3) \quad . \tag{2.10}$$

In der hier vorliegenden Form sind die dielektrische Funktion bzw. ihre Inverse nicht symmetrisch bezüglich der Variablen (1) und (2). Die daraus resultierenden numerischen Probleme, insbesondere die auftretenden Divergenzen der räumlichen Fouriertransformierten im Fall $\mathbf{q}\rightarrow\mathbf{0}$, umgehen wir, indem wir mit symmetrisierten dielektrischen Matrizen arbeiten (siehe Abschnitt 3.3 und Anhang D).

Sämtliche Erweiterungen und Ergänzungen, die über die reine GW-Näherung für Σ hinausgehen, werden in der Literatur als *Vertex-Korrekturen* bezeichnet. Darunter finden sich zum einen Ansätze, in denen die Entwicklung der Selbstenergie nach der Wechselwirkung W partiell bis in unendliche Ordnung berücksichtigt wird [Nor87, Nor89, Mah89, Sol94]. Das kann etwa durch Einbeziehung einer lokalen Austausch-Korrelations-Wechselwirkung $v_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = \delta^2 E_{\mathrm{xc}}/\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')$ in die Wechselwirkung der Elektronen geschehen. [E_{xc} ist das Austausch-Korrelations-Energiefunktional der DFT.] Andere Arbeiten zielen darauf ab, in sytematischer Weise genau eine höhere Ordnung (d.h. die zweite Ordnung) der Reihenentwicklung auszuwerten und danach die Reihe wiederum abbrechen zu lassen [Bob94]. Welche Vorgehensweise am ehesten geeignet ist, den Anwendungsbereich der GW-Näherung signifikant zu erweitern, ist nach wie vor ungeklärt und umstritten. Viele Arbeiten scheinen darauf hin zu deuten, daß der Beitrag der Vertex-Korrekturen zur Valenz- und Leitungsbandstruktur von einfachen Metallen und von typischen Volumen-Halbleitern (etwa Si) marginal ist und für die meisten Fragestellungen vernachlässigt werden kann [Mah89, Sol94, Bob94]. Bei stärker korrelierten Elektronensystemen, wie etwa den Übergangsmetallen, stellt sich hingegen heraus, daß wesentliche Aspekte der elektronischen Struktur vom reinen GW-Zugang nicht erfaßt werden können. Beispielsweise finden sich im Photoemissionsspektrum von Ni und von NiO zusätzlich zu den Quasiteilchen weitere Satellitenstrukturen, die durch die gleichzeitige Anregung zweier d-Elektronen am gleichen Ni-Atom hervorgerufen werden. Die zugehörigen Onsite-Korrelationseffekte zwischen Elektronen an einem einzigen Ni-Atom können durch die reine GW-Näherung prinzipiell nicht beschrieben werden [Ary92, Ary95]. Die

Verläßlichkeit der GW-Näherung bei der Anwendung auf stark korrelierte Systeme wurde an Hubbard-Modellsystemen untersucht [Ver95, Pol95]. Hier können die Resultate der GWA direkt mit exakten Ergebnissen verglichen werden. Es stellt sich heraus, daß in der Spektralfunktion deutliche Abweichungen zwischen dem GW-Ergebnis und dem exakten Resultat bestehen. Diese Schwächen der GWA lassen sich nur durch Hinzunahme höherer Ordnungen in der Reihenentwicklung beheben. Für schwach korrelierte Systeme hingegen sind Vertex-Korrekturen grundsätzlich von geringer Bedeutung. Wir werden uns daher in dieser Arbeit auf die reine GW-Näherung konzentrieren und keine Vertex-Korrekturen behandeln.

2.2 Quasiteilchen-Korrekturen

Wenden wir uns nun der Frage zu, wie die GW-Näherung konkret ausgewertet werden kann. Die selbstkonsistente Lösung des Gleichungssystems (2.5, 2.6) könnte zum Beispiel in iterativer Weise mit einem ständig zu aktualisierenden Selbstenergie-Operator $\Sigma[G]$ geschehen. Dies erweist sich, insbesondere wegen der siebendimensionalen Abhängigkeit der Green-Funktion G von $\mathbf{r}$, $\mathbf{r}'$ und $(t - t')$, als nahezu undurchführbar. In dieser Richtung sind bisher nur erste Ansätze unternommen worden [Roj95]. Statt dessen verfolgen wir in dieser Arbeit einen anderen Zugang. Er beruht darauf, daß zur Beschreibung von Einteilchen-Prozessen hauptsächlich die elektronischen Quasiteilchen von Interesse sind. Die Quasiteilchen $|\mathrm{m}\mathbf{k}\rangle$ sind als scharfe Peaks in der Spektralfunktion deutlich zu erkennen [Roj95]. Sie entsprechen langlebigen Anregungen des Systems durch Hinzufügen ($E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} > E_F$) bzw. Entfernen ($E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} < E_F$) eines Elektrons und besitzen daher Teilchencharakter. Wegen der Wechselwirkung innerhalb des Systems kann dieser Teilchencharakter aber nicht einem einzelnen Elektron zugeordnet werden. Die Quasiteilchen sind durch Quasiteilchen-Amplituden $\psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}(\mathbf{x})$ und -Energien $E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}$ charakterisiert, die sich als verallgemeinerte Eigenvektoren und -werte aus Gl. (2.5) ergeben:

$$\begin{aligned} \left\{-\nabla^2 + V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{Coul}}(\mathbf{r})\right\} \psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}(\mathbf{x}) & \\ + \int \Sigma(\mathbf{x}, \mathbf{x}', E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) \psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}(\mathbf{x}') d\mathbf{x}' &= E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} \psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}(\mathbf{x}) \quad . \end{aligned} \tag{2.11}$$

Der Selbstenergie-Operator ist im allgemeinen nicht hermitesch. Dementsprechend bilden die Quasiteilchen-Amplituden kein orthonormales Funktionensystem, und die Quasiteilchen-Energien können komplex sein. Der Kehrwert ihres Imaginärteils läßt

sich als Lebensdauer der elementaren Anregung interpretieren. Die Lebensdauer ist prinzipiell endlich, denn wegen der Wechselwirkung können die Quasiteilchen mittels Streuung am Elektronensystem in andere Quasiteilchen zerfallen. Wir beschränken uns jedoch in unseren Rechnungen zunächst auf rein reelle Energien.

Auch im Rahmen dieses eingeschränkten Quasiteilchen-Konzeptes erweist sich die Auswertung der GWA als so aufwendig, daß eine selbstkonsistente Behandlung von (2.5) bzw. (2.11) für die meisten Systeme nicht durchgeführt werden kann. Das ist in der Regel auch gar nicht notwendig, sofern Startwerte ψ^{Start} und E^{Start} zur Verfügung stehen, die den zu berechnenden Größen ψ^{QP} und E^{QP} ausreichend nahe kommen. Diese Forderung wird in hervorragender Weise von den Ergebnissen einer vorangehenden DFT-LDA-Rechnung erfüllt. Bei vielen Halbleitersystemen stellt sich heraus, daß die Übereinstimmung der (nicht renormierten) Quasiteilchen-Amplituden ψ^{QP} mit den LDA-Wellenfunktionen ψ^{LDA} besser als 99% ist [Hyb86, God88], und daß die Kohn-Sham-Energien der LDA von den Quasiteilchen-Energien um weniger als 10% der Plasmafrequenz des betrachteten Systems voneinander abweichen. Daraus folgen zwei wesentliche Punkte:

- Der Selbstenergie-Operator kann in guter Näherung aus den DFT-LDA-Ergebnissen konstruiert werden.
- Aus dem Vergleich

$$\begin{aligned} & \{-\nabla^2 + V_{\mathrm{ext}} + V_{\mathrm{Coul}}\}\,\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} \;+\; V^{\mathrm{LDA}}_{\mathrm{xc}}\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} \;=\; E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} \\ \text{und} \quad & \{-\nabla^2 + V_{\mathrm{ext}} + V_{\mathrm{Coul}}\}\,\psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} \;+\; \Sigma(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}})\psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} \;=\; E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}\psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} \end{aligned}$$

  folgt unter der Annahme $\psi^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} \approx \psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}$ die Beziehung

$$E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} + \left\langle \psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} | \Sigma(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) - V^{\mathrm{LDA}}_{\mathrm{xc}} | \psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} \right\rangle \quad . \tag{2.12}$$

Der Beitrag auf der rechten Seite von Gl. (2.12), der zu der LDA-Energie hinzuaddiert wird, wird allgemein als *Quasiteilchen-Korrektur* zur LDA bezeichnet. Formal entspricht Gl. (2.12) einer Störungsrechnung in erster Ordnung für die Energie. Der Störoperator ist durch die Differenz $(\Sigma - V^{\mathrm{LDA}}_{\mathrm{xc}})$ gegeben. Das suggeriert einen engen Zusammenhang zwischen der LDA und der GW-Näherung. Die LDA tritt hier als „Vorstufe“ der GW-Näherung auf (etwa bei der Konstruktion des Selbstenergie-Operators aus DFT-Größen oder bei der Berechnung der Quasiteilchen-*Korrekturen*

zu LDA-Bandstrukturen). Außerdem werden häufig GWA-Resultate mit LDA-Ergebnissen verglichen. Diese Verknüpfung entsteht einerseits aus der historischen Entwicklung der Methoden, denn erst nach dem Scheitern der LDA bei der Berechnung von elektronischen Übergangsenergien wurde die GW-Näherung auf Halbleiter angewendet. Andererseits ist eine numerische Ähnlichkeit zwischen den beiden Zugängen unverkennbar vorhanden, am deutlichsten bei der oft sehr guten Übereinstimmung der DFT-Wellenfunktionen mit den (nicht renormierten) Quasiteilchen-Amplituden. Bezüglich der theoretischen Grundlagen besteht jedoch kein direkter systematischer Zusammenhang zwischen den beiden Ansätzen. Es handelt sich um völlig verschiedene Methoden, die klar voneinander abzugrenzen sind. Das äußert sich unter anderem darin, daß für den Störoperator $(\Sigma - V_{\mathrm{xc}}^{\mathrm{LDA}})$ kein geschlossener Ausdruck existiert. Vielmehr muß jeder der beiden Terme separat bestimmt werden, bevor ihre Differenz gebildet werden kann.

Zur Verwendung der Begriffe DFT und LDA merken wir folgendes an. Vom Standpunkt dieser Arbeit aus betrachtet, stellt die LDA eine Möglichkeit zur Realisierung der DFT bei der Anwendung auf ein konkretes System dar. Ob die für die GWA notwendigen Größen in der LDA oder in der exakten DFT bestimmt werden, ist wegen der numerischen Ähnlichkeit zwischen DFT und LDA [God88] für uns ohne Belang. Beide Begriffe werden daher von uns weitgehend synonym verwendet; für eine nachfolgende GW-Rechnung ergeben sich keine grundsätzlich neuen Erkenntnisse, wenn sie auf einer exakten DFT-Rechnung anstelle einer LDA-Rechnung aufbaut. Wir sprechen dennoch hauptsächlich von LDA-Rechnungen, um nicht den Eindruck entstehen zu lassen, wir könnten exakte DFT-Rechnungen durchführen.

Analog zu Gl. (2.12) lassen sich auch Störungsbeiträge zweiter Ordnung zur Energie bestimmen, die mit Störungsbeiträgen erster Ordnung zu den Quasiteilchen-Amplituden zusammenhängen. Diese Störungsbeiträge zweiter Ordnung zur Energie betragen bei den meisten von uns untersuchten Materialien weniger als 0.05 eV. Daraus leitet sich letztlich die oben erwähnte sehr gute Übereinstimmung der LDA-Wellenfunktionen mit den Quasiteilchen-Amplituden ab. Bei einigen Systemen treten allerdings auch deutliche Abweichungen auf, wie etwa bei Ge, wenn es mit einem Pseudopotential beschrieben wird, das die Elektronen der dritten Schale als Valenzelektronen berücksichtigt (siehe Abschnitt 13.4). In solchen Fällen bedarf es einer genaueren Behandlung des Störoperators, zum Beispiel einer expliziten Neuberechnung seiner Eigenvektoren und -werte innerhalb der GWA.

Die oftmals sehr gute Übereinstimmung der Wellenfunktionen ist bisher nicht

vollständig verstanden. Sie deutet darauf hin, daß das Austausch-Korrelations-Potential V_{xc} bereits eine erstaunlich gute Approximation des Selbstenergie-Operators Σ darstellt. Man sollte jedoch auch bedenken, daß die Konstruktion des Selbstenergie-Operators aus LDA-Wellenfunktionen eine wesentliche Einschränkung der Physik eines Systems bedeuten könnte, die möglicherweise durch die Bestimmung der Quasiteilchen-Amplituden aus diesem eingeschränkten Σ gar nicht überwunden werden kann. Konkrete Beispiele für solche Probleme sind uns allerdings nicht bekannt.

Mit den oben angeführten Überlegungen ist nun die durchzuführende Arbeitsweise wie folgt festgelegt.

1. Ermittlung der atomaren Struktur des zu behandelnden Systems (experimentell oder theoretisch, etwa im Rahmen der DFT-LDA).

2. Berechnung der Wellenfunktionen $\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}(\mathbf{r})$ und der zugehörigen Energien $E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}$ im Rahmen der LDA.

3. Konstruktion des Selbstenergie-Operators Σ aus den Größen $\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}$ und $E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}$.

4. Auswertung der Matrix des Störoperators $\Sigma - V_{\mathrm{xc}}$ in den LDA-Wellenfunktionen:
$$\left\langle \psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} | \Sigma(E) - V_{\mathrm{xc}} | \psi^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}} \right\rangle \tag{2.13}$$

Der letzte Punkt verdient noch eine genauere Diskussion. Insbesondere müssen wir die Energie-Abhängigkeit des Selbstenergie-Operators $\Sigma(E)$ noch untersuchen. Aus den Diagonalelementen der Matrix $(\Sigma(E^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}}) - V_{\mathrm{xc}})_{\mathrm{mm}'}$ (also mit Einteilchen-Energien, die durch die LDA gegeben sind) erhält man die Störungsbeiträge erster Ordnung:
$$\Delta^{\mathrm{QP},1}_{\mathrm{m}\mathbf{k}} = (\Sigma(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}})_{\mathrm{mm}} \quad . \tag{2.14}$$
Die Nichtdiagonalelemente ergeben Beiträge in zweiter Ordnung
$$\Delta^{\mathrm{QP},2}_{\mathrm{m}\mathbf{k}} = \sum_{\mathrm{m}' \neq \mathrm{m}} \frac{|(\Sigma(E^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}}) - V_{\mathrm{xc}})_{\mathrm{mm}'}|^2}{E^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}} - E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}} \quad , \tag{2.15}$$

Wir merken an, daß zur Bestimmung der Beiträge zweiter Ordnung bei entarteten Zuständen keine zusätzliche Diagonalisierung des Störoperators im entarteten Unterraum notwendig wird, denn der Selbstenergie-Operator besitzt dieselben Symmetrie-Eigenschaften wie das Austausch-Korrelations-Potential und hebt die Entartung

nicht auf. Die Nichtdiagonalelemente und somit auch die Beiträge zweiter Ordnung zur Energie stellen sich in den meisten Fällen als sehr gering heraus. Falls die Nichtdiagonalelemente groß sein sollten, so empfiehlt es sich, den störungstheoretischen Zugang zu verlassen und statt dessen in einem geeigneten Unterraum {m} (aufgespannt durch die entsprechenden LDA-Wellenfunktionen) den Hamilton-Operator

$$\left(-\nabla^2 + V_{\text{ext}} + V_{\text{Coul}} + \Sigma(E)\right)_{\text{mm}'}(\mathbf{k}) \equiv E^{\text{LDA}}_{\text{m}'\mathbf{k}}\delta_{\text{mm}'} + (\Sigma(E) - V_{\text{xc}})_{\text{mm}'} \tag{2.16}$$

explizit zu diagonalisieren (siehe Abschnitt 13.4).

Auch wenn die Terme $\Delta^{\text{QP},2}_{\text{m}\mathbf{k}}$ verschwinden, so ergeben sich noch zusätzliche Energiebeiträge aus der expliziten Energieabhängigkeit des Selbstenergie-Operators $\Sigma(E)$. Diese Energieabhängigkeit verhält sich in erster Näherung linear (vgl. Abb. 5.2):

$$\Sigma_{\text{mm}'}(E) = \Sigma_{\text{mm}'}(\bar{E}) + \Sigma'_{\text{mm}'}(\bar{E}) \cdot (E - \bar{E}) \tag{2.17}$$

mit einer Ableitung Σ', die über einen weiten Energiebereich um die Quasiteilchen-Energie herum kaum variiert. In einer systematischen Störungsentwicklung liefert diese Energieabhängigkeit in n-ter Ordnung einen Energiebeitrag

$$[\Sigma'_{\text{mm}}]^{n-1} \cdot \Delta^{\text{QP},1}_{\text{m}\mathbf{k}} \tag{2.18}$$

zur Quasiteilchen-Korrektur. Als geometrische Reihe $\sum_{n=0}^{\infty} x^n = 1/(1-x)$ ergibt sich letztlich für die Energieänderung $\Delta^{\text{QP}}_{\text{m}\mathbf{k}}$:

$$\Delta^{\text{QP}}_{\text{m}\mathbf{k}} = Z_{\text{m}\mathbf{k}} \cdot \Delta^{\text{QP},1}_{\text{m}\mathbf{k}} + \Delta^{\text{QP},2}_{\text{m}\mathbf{k}} + \ldots \tag{2.19}$$

mit der *Renormierungs-Konstanten*

$$Z_{\text{m}\mathbf{k}} = \frac{1}{1 - \Sigma'_{\text{mm}}} \quad . \tag{2.20}$$

Die Renormierungs-Konstante $Z_{\text{m}\mathbf{k}}$ besitzt eine physikalische Bedeutung, die am leichtesten am homogenen Elektronengas deutlich wird. Für Zustände an der Fermikante gibt sie die Unstetigkeit der Besetzungsfunktion an: $Z_{k_{\text{F}}} = f_{k_{\text{F}}^-} - f_{k_{\text{F}}^+}$ [Ful91]. Da die Besetzungsfunktion f_k wegen des Pauli-Prinzips zwischen 0 und 1 liegen muß und f_k mit k monoton fällt, folgt

$$0 < Z_{k_{\text{F}}} < 1 \tag{2.21}$$

$$\text{und somit} \qquad \frac{\partial \Sigma_k(E)}{\partial E} < 0 \quad . \tag{2.22}$$

Die Selbstenergie fällt also monoton (vergleiche Abschnitt 5). Für die Renormierungs-Konstante ergeben sich in unseren Rechnungen Werte, die im allgemeinen zwischen 0.6 und 0.8 liegen. Für ein homogenes Elektronengas mit einer entsprechenden Elektronendichte erhält man ähnliche Ergebnisse [Hed69].

Die Renormierungs-Konstante $Z_{\mathrm{m}\mathbf{k}}$ entspricht dem spektralen Gewicht des zugehörigen Quasiteilchen-Peaks im elektronischen Spektrum des Zustands $|\mathrm{m}\mathbf{k}\rangle$. Da das gesamte spektrale Gewicht eines Zustands normiert sein muß, folgt aus $Z_{\mathrm{m}\mathbf{k}} < 1$ unmittelbar, daß zusätzlich zu den Quasiteilchen noch weitere Anregungen des Systems möglich sind. Neben den Quasiteilchen-Peaks treten im Spektrum zusätzliche *Satelliten-Strukturen* in Erscheinung, deren Energien um 1–2 Plasmafrequenzen unterhalb oder oberhalb der Quasiteilchen-Energie liegen [Bec94, Roj95]. Sie werden bei Hedin und Lundqvist als *Plasmaronen* bezeichnet [Hed69]. Hieraus ergibt sich sofort, daß die Quasiteilchen mit ihren Amplituden und Energien nicht die vollständige Information über das elektronische System beinhalten. Folglich läßt sich die Green-Funktion aus den Quasiteilchen allein nicht exakt konstruieren. In Kap. 14 werden wir diskutieren, wie man zu einem vollständigeren Bild gelangen kann, das besonders auf die Satellitenstruktur eingeht.

Bei der konkreten Ausführung von Rechnungen innerhalb des GW-Formalismus stellt sich die zentrale Frage nach der Darstellung der von $(\mathbf{r}, \mathbf{r}', t)$ abhängigen Funktionen P, W, ϵ und ϵ^{-1}. Die scheinbar einfachste Lösungsmöglichkeit besteht in der Einführung äquidistanter Netze in Raum und Zeit. Dieser Zugang erweist sich jedoch als extrem aufwendig und konnte bisher nur auf einfache Systeme angewendet werden [Roj95]. Um GW-Rechnungen auf kleinen Rechnersystemen in kurzer Zeit durchzuführen, und um den Zugang auf komplexere Systeme anwenden zu können, ist eine Darstellung der Funktionen mit Hilfe von Basisfunktionen notwendig.

Für die räumliche Abhängigkeit von $\mathbf{r}$ und $\mathbf{r}'$ wird von den meisten Autoren eine Fouriertransformation durchgeführt, d.h. alle Funktionen werden mittels ebener Wellen dargestellt. Als Alternative dazu verwenden wir in dieser Arbeit lokalisierte Gauß-Orbitale als Basissatz. Dieser Zugang orientiert sich an der räumlich inhomogenen Struktur halbleitender Elektronensysteme. Die Inhomogenität führt zu entsprechenden räumlichen Eigenschaften der Green-Funktion und aller anderen Größen, die im GW-Zugang auftreten. Eine darauf abgestimmte Wahl räumlicher Ansatzfunktionen erweist sich besonders bei Oberflächensystemen und bei der Behandlung von Semicore-Zuständen als extrem nützlich, denn der numerische Aufwand wird im Vergleich zum Ebenen-Wellen-Zugang enorm reduziert. Das betrifft sowohl die er-

forderlichen Rechenzeiten als auch den benötigten Hauptspeicherbedarf; hierdurch wird eine Behandlung der vorgestellten Systeme mit den uns zur Verfügung stehenden Rechnersystemen überhaupt erst möglich. Die Behandlung der räumlichen Abhängigkeit wird im nächsten Kapitel diskutiert.

Der zeitliche Verlauf der abgeschirmten Wechselwirkung ist besonders durch das Auftreten oszillierender Plasmonen gekennzeichnet. Es stellt sich heraus, daß die für die Berechnung von Quasiteilchen relevanten Eigenschaften dieser Dynamik durch entsprechende Plasmon-Pol-Modelle sehr gut erfaßt werden. Diese Modelle erläutern wir in Kap. 4.

Kapitel 3

Darstellung von Zweipunkt-Funktionen

In diesem Kapitel behandeln wir die Darstellung von Zweipunkt-Funktionen $f(\mathbf{r}, \mathbf{r}')$ mittels geeigneter Basisfunktionen. Die zusätzliche Abhängigkeit der Funktionen von der Zeit t wird hier zunächst nicht betrachtet.

3.1 Ebene Wellen und Gauß-Orbitale

In der Literatur bedient sich die überwiegende Zahl von Arbeiten zur GWA der Darstellung der Zweipunkt-Funktionen P, ϵ, ϵ^{-1} und W als Fourierreihe, also mit Hilfe ebener Wellen (plane waves, PW) als Basisfunktionen. Diese Vorgehensweise, die wir im folgenden als *PW-GWA* bezeichnen, weist zwei wichtige Eigenschaften auf. Zum einen erhalten in dieser Basis alle Rechenschritte wegen der Orthonormalität der ebenen Wellen eine besonders einfache Struktur und sind somit programmtechnisch leicht zu realisieren. Zum anderen kann durch Heraufsetzen der Abschneideenergie in systematischer Weise Konvergenz bezüglich des Basissatzes erzielt werden. Es gibt jedoch zwei Klassen von Systemen, bei denen sich ebene Wellen als Basisfunktionen als weit weniger geeignet herausstellen als etwa die von uns verwendeten Gauß-Orbitale (GO):

- Zur Untersuchung von *Systemen mit reduzierter Symmetrie* (Oberflächen, Grenzflächen, Defekte, Legierungen, Quantenstrukturen uva.) sind große Superzellen mit vielen Atomen notwendig. Wenn n die Anzahl der Atome in der Superzelle ist, und wenn zur Erfassung der elektronischen Größen m Basis-

funktionen pro Atom erforderlich sind, so skalieren Speicherbedarf und Rechenzeit für die GWA im wesentlichen wie m^2n^2 oder m^2n^3 (je nach Art des verwendeten Plasmon-Pol-Modells). Bei den von uns untersuchten Halbleitermaterialien stellt sich heraus, daß ein GO-Basissatz erheblich weniger Ansatzfunktionen *m pro Atom* zu einer konvergenten Beschreibung erfordert, als dies bei ebenen Wellen der Fall wäre. Daraus eröffnet sich die Möglichkeit, durch Verwendung einer GO-Basis den Rechenaufwand zu reduzieren. Diese Rechenmethode, die wir hier vorstellen, wird im weiteren als *GO-GWA* bezeichnet. Wegen der Nicht-Orthogonalität der Gauß-Orbitale ist die Struktur der Rechenschritte komplizierter als bei der PW-GWA. Daher gewinnen wir bei der Untersuchung herkömmlicher Volumen-Halbleiter keine Vorteile gegenüber der PW-GWA. Die Behandlung von Oberflächen im Rahmen einer Superzellen-Rechnung wird hingegen durch die Verwendung von GO's um ein Mehrfaches beschleunigt und läßt sich auf relativ kleinen Rechnern durchführen. Der Rechenzeitgewinn wird noch dadurch verstärkt, daß eine lokalisierte Basis gegenüber einer delokalisierten Basis (etwa aus ebenen Wellen) zu einer Verbesserung der Skalierung mit der Systemgröße n führt. Im Extremfall einer echten Ortsdarstellung beträgt der Gewinn eine Potenz in n [Roj95]; dies wird von unserer GO-GWA nicht ganz erreicht.

- Eine völlig andere Situation ist gegeben, wenn *Core- und Semicore-Zustände in Halbleitern* explizit berücksichtigt oder selber untersucht werden sollen. Hier spielt die Systemgröße n eine untergeordnete Rolle. Schon bei Volumensystemen tritt die entscheidende Schwierigkeit auf, daß die Core- und Semicore-Elektronen im Vergleich zur Ausdehnung der Valenzelektronen und zum atomaren Bindungsabstand stark lokalisiert sind. In PW-Rechnungen steigt die notwendige Zahl m an ebenen Wellen pro Atom um Größenordnungen an, verglichen mit einer Untersuchung, in der nur die Valenzelektronen berücksichtigt werden. Schon LDA-Rechnungen stellen einen enormen Aufwand dar, und GW-Rechnungen mit ebenen Wellen sind für solche Fragestellungen bis auf wenige Ausnahmen nahezu undurchführbar. Hier schaffen lokalisierte Ansatzfunktionen, z. B. Gauß-Orbitale, in einfacher Weise Abhilfe. Die Wellenfunktionen der Core-Zustände sind an den Positionen der Atomkerne zentriert. Sie lassen sich daher durch Orbital-Ansatzfunktionen, die ebenfalls an den Kernen zentriert werden, sehr effizient beschreiben. Die starke Lokalisie-

rung der Core- und Semicore-Elektronen an den Atomkernen läßt sich durch Hinzunahme von wenigen, entsprechend steilen Radialfunktionen problemlos erfassen. Schon bei der LDA-Behandlung von II-VI-Halbleitern unter Berücksichtigung der kationischen d-Zustände wurde der enorme Einspareffekt der GO gegenüber einer PW-Basis deutlich [Sch93a, Sch93b]. Bei den in unserer Arbeit durchgeführten GO-GWA-Untersuchungen solcher Systeme stellen wir eine Reduktion von Speicherbedarf und Rechenzeit um eine Größenordnung und mehr gegenüber der PW-GWA fest. Dadurch werden diese Rechnungen auf den uns zur Verfügung stehenden Rechnersystemen überhaupt erst möglich.

Der Übersichtlichkeit halber beginnen wir die Diskussion mit der Darstellung der Zweipunkt-Funktionen als Fourierreihe, also mit Hilfe ebener Wellen als Basisfunktionen. In den von uns behandelten kristallinen Systemen liegt ein gitterperiodisches Verhalten von $f(\mathbf{r}, \mathbf{r}')$ vor:

$$f(\mathbf{r}, \mathbf{r}') = f(\mathbf{r} + \mathbf{R}, \mathbf{r}' + \mathbf{R}) \quad , \tag{3.1}$$

$\mathbf{R}$ eine beliebige Gittertranslation.

Daraus ergibt sich die folgende Darstellung von f als Fourier-Reihe:

$$f(\mathbf{r}, \mathbf{r}') = \frac{1}{V} \sum_{\mathbf{q}} \sum_{\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}} f_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) e^{-i(\mathbf{q}+\mathbf{G}')\mathbf{r}'} \tag{3.2}$$

($\mathbf{G}, \mathbf{G}'$ rez. Gittervektoren; $\quad \mathbf{q} \in 1.\mathrm{BZ}$) .

[V ist das Volumen der Einheitszelle.] Die Matrixelemente $f_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ berechnen sich als

$$f_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \frac{1}{V} \int e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}} f(\mathbf{r}, \mathbf{r}') e^{i(\mathbf{q}+\mathbf{G}')\mathbf{r}'} d^3r d^3r' \quad . \tag{3.3}$$

Prinzipiell wird die Darstellung durch die ebenen Wellen $e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}$ in (3.2) dadurch möglich, daß die Funktionen $\{e^{i\mathbf{G}\mathbf{r}}\}$ eine vollständige Basis für den Raum der gitterperiodischen Einpunkt-Funktionen bilden. Diese Eigenschaft erfüllen natürlich auch andere Funktionensysteme. Wir benutzen in dieser Arbeit gitterperiodische Superpositionen von Gauß-Orbitalen $\phi_\beta(\mathbf{r})$, die an den atomaren Positionen lokalisiert sind. Lokalisierte Ansatzfunktionen, insbesondere Gauß-Orbitale, werden seit langem erfolgreich in der Quantenchemie zur Beschreibung von Molekülen verwendet [Oba85]. In Form der gitterperiodischen Superpositionen werden Gauß-Orbitale mit großem Erfolg in LDA-Untersuchungen von Halbleitern und Halbleiteroberflächen

zur Darstellung der Wellenfunktionen verwendet. Sie weisen eine genügend hohe Flexibilität zur Darstellung der physikalisch relevanten Funktionen auf. Man kann sie auch als geeignete Linearkombinationen von ebenen Wellen verstehen, nach denen sie sich ihrerseits entwickeln lassen. Wir gewinnen in völliger Analogie zur Fourier-Reihe (3.2) eine Darstellung der Zweipunkt-Funktionen in der Form

$$f(\mathbf{r},\mathbf{r}') = \sum_{\mathbf{q}} \sum_{\beta\beta'} \chi_\beta(\mathbf{q},\mathbf{r}) \langle f \rangle_{\beta\beta'}(\mathbf{q}) \chi^*_{\beta'}(\mathbf{q},\mathbf{r}') \tag{3.4}$$

mit den $\mathbf{q}$-abhängigen Linearkombinationen der Gauß-Orbitale:

$$\chi_\beta(\mathbf{q},\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{q}(\mathbf{R}+\boldsymbol{\tau}_\beta)} \phi_\beta(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_\beta) \quad . \tag{3.5}$$

[N ist die Gesamtzahl der Einheitszellen.] Die Gauß-Orbitale ϕ_β sind durch ihren Aufpunkt $\boldsymbol{\tau_\beta}$ (Position eines Atomkerns), ihre Abklingkonstante β und ihren Orbitaltyp (s, p_x, ..., d_{xy} etc.) gekennzeichnet. Den Orbitaltyp bezeichnen wir nicht durch einen eigenen Index; im allgemeinen gehört zu einer Abklingkonstante β ein Gesamtpaket aus verschiedenen Orbitaltypen. [Die Ansatzfunktionen sind von genau der gleichen Art, wie wir sie auch bei den LDA-Rechnungen zur Darstellung der Wellenfunktionen verwenden (siehe Gl. (3.10)). Allerdings können die Abklingkonstanten und Orbitaltypen β unabhängig von der LDA-Basis $\{\alpha\}$ gewählt werden.] Analog zur Fouriertransformation erfolgt auch hier die Darstellung der Zweipunkt-Funktionen mit Hilfe von Matrizen in β und β', die zusätzlich noch vom Wellenvektor $\mathbf{q}$ aus der ersten Brillouin-Zone abhängen. Allerdings treten zwei verschiedene Typen solcher Matrizen in Erscheinung, die wir in dieser Arbeit mit $[f]_{\beta\beta'}(\mathbf{q})$ und $\langle f \rangle_{\beta\beta'}(\mathbf{q})$ bezeichnen. Aus der Matrix $\langle f \rangle$ läßt sich gemäß Gl. (3.4) die Funktion $f(\mathbf{r},\mathbf{r}')$ bestimmen. Ihre Berechnung aus $f(\mathbf{r},\mathbf{r}')$ muß jedoch in zwei Schritten erfolgen. Zunächst werden die Matrixelemente $[f]_{\beta\beta'}(\mathbf{q})$ wie folgt bestimmt:

$$[f]_{\beta\beta'}(\mathbf{q}) = \int \chi^*_\beta(\mathbf{q},\mathbf{r}) f(\mathbf{r},\mathbf{r}') \chi_{\beta'}(\mathbf{q},\mathbf{r}') d^3r d^3r' \quad . \tag{3.6}$$

Diese Matrix $[f]$ läßt sich dann in die Matrix $\langle f \rangle$ umrechnen:

$$\langle f \rangle = S^{-1}[f]S^{-1} \quad , \tag{3.7}$$

wobei die Inverse der Überlappmatrix

$$S_{\beta\beta'}(\mathbf{q}) = \int \chi^*_\beta(\mathbf{q},\mathbf{r}) \chi_{\beta'}(\mathbf{q},\mathbf{r}) d^3r \tag{3.8}$$

auftritt. Zur ausführlichen Diskussion der Eigenschaften dieser Transformationen verweisen wir auf Anhang A.

Bei der Fourier-Darstellung mittels ebener Wellen ist die Überlapp-Matrix durch die Einheitsmatrix gegeben ($S=\underline{\underline{1}}$), da die ebenen Wellen ein orthonormiertes System darstellen, und zwischen den entsprechenden Matrizen $\langle f\rangle$ und $[f]$ braucht gar nicht unterschieden zu werden. Die von uns verwendeten Gauß-Orbitale sind hingegen nicht orthonormiert, so daß $S\neq\underline{\underline{1}}$ gilt. Das erfordert die Einführung der zwei verschiedenen Matrizen $[f]_{\beta\beta'}(\mathbf{q})$ und $\langle f\rangle_{\beta\beta'}(\mathbf{q})$.

3.2 Polarisierbarkeit

Wenden wir uns nun der Realisierung der GWA mit Hilfe dieser Darstellungstechnik zu. Zunächst ist die Polarisierbarkeit $P=-iGG$ zu bestimmen. Die Green-Funktion G konstruieren wir aus den Ergebnisse der vorangehenden LDA-Rechnung zu

$$G(\mathbf{x},\mathbf{x}',E)=\sum_{\mathrm{m}\mathbf{k}}\frac{\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}(\mathbf{x})\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}(\mathbf{x}')^*}{E-E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}+i0^+\mathrm{sign}(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}-E_F)} \quad . \tag{3.9}$$

Bei den in dieser Arbeit untersuchten halbleitenden Systemen nehmen die Elektronen $|\mathrm{m}\mathbf{k},\sigma{=}{+}1\rangle$ und $|\mathrm{m}\mathbf{k},\sigma{=}{-}1\rangle$ die gleichen Wellenfunktionen $\psi_{\mathrm{m}\mathbf{k}}(\mathbf{r})$ und die gleichen Energien $E_{\mathrm{m}\mathbf{k}}$ an; zwischen Minoritäts- und Majoritäts-Ladungsträgern braucht nicht unterschieden zu werden. Wir notieren daher die Spinkoordinate σ im folgenden nicht mehr explizit. An einigen Stellen in der GWA sind die Spinentartung sowie die Orthogonalität von Zuständen verschiedenen Spins zu berücksichtigen. Die Wellenfunktionen $\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}(\mathbf{r})$ stellen wir als Linearkombinationen von Gauß-Orbitalen dar:

$$\psi^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}(\mathbf{r}) \;=\; \sum_{\alpha} C^{\alpha}_{\mathrm{m}\mathbf{k}}\chi_{\alpha}(\mathbf{k},\mathbf{r}) \tag{3.10}$$

$$\text{mit}\quad \chi_{\alpha}(\mathbf{k},\mathbf{r}) \;=\; \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} e^{i\mathbf{k}(\mathbf{R}+\boldsymbol{\tau}_{\alpha})}\phi_{\alpha}(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_{\alpha}) \quad . \tag{3.11}$$

Die Gauß-Orbitale ϕ_α für die LDA-Basis sind von der gleichen Struktur wie die Gauß-Orbitale ϕ_β der GW-Basis. Aus (3.9) berechnet sich die Darstellungsmatrix der Polarisierbarkeit zu

$$P_{\beta\beta'}(\mathbf{q},\omega)=2\frac{1}{V}\sum_{\mathbf{k}}\sum_{\mathrm{m}\in\mathrm{Val}}\sum_{\mathrm{n}\in\mathrm{Con}} M^{\mathrm{mn}}_{\beta}(\mathbf{k},\mathbf{q})\left[M^{\mathrm{mn}}_{\beta'}(\mathbf{k},\mathbf{q})\right]^*$$

$$\times\quad\left[\frac{1}{E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}-E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}-\omega+i0^+}+\frac{1}{E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}-E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}+\omega+i0^+}\right] \quad . \tag{3.12}$$

Der Faktor 2 trägt der Spinentartung Rechnung. Es wurde die Abkürzung

$$M_\beta^{\mathrm{mn}}(\mathbf{k},\mathbf{q}) = \int \psi_{\mathrm{m}\mathbf{k}}^{\mathrm{LDA}*}(\mathbf{r})\chi_\beta^*(\mathbf{q},\mathbf{r})\psi_{\mathrm{n},\mathbf{k}+\mathbf{q}}^{\mathrm{LDA}}(\mathbf{r})d^3r \tag{3.13}$$

verwendet. An dieser Stelle wird deutlich, wie die Basis $\{\chi_\beta\}$ zu wählen ist: Die Basisfunktionen müssen flexibel genug sein, um *Produkte zweier Wellenfunktionen* darzustellen. Betrachten wir etwa diejenigen Wellenfunktionsanteile mit dem größten Drehimpuls. Das entsprechende Produkt der Wellenfunktion mit sich selber wird dementsprechend den doppelten Drehimpuls beinhalten. Daher muß der GW-Basissatz $\{\chi_\beta\}$ höhere Drehimpulskomponenten als der LDA-Basissatz beinhalten. Im Unterschied zu atomaren Systemen kann diese Analyse für Festkörper nicht streng durchgeführt werden, da die Orbitale eines Atoms, von einem anderen Atom aus betrachtet, Beiträge zu jedem Drehimpuls liefern. Wenn ebene Wellen als GW-Basissatz benutzt werden, so stellt sich die Frage nach den Drehimpulsanteilen gar nicht, denn ebene Wellen beinhalten von sich aus sämtliche Drehimpulskomponenten.

Als LDA-Basissatz verwenden wir Orbitale mit den Drehimpulsen l=0, 1 und 2. Demzufolge erfordert der GW-Basissatz Orbitale mit l=0–4, also s-, p-, d-, f- und g-Orbitale. In vielen Systemen kann auf die g-Orbitale verzichtet werden. Die Integrale (3.13) berechnen sich als

$$\begin{aligned} M_\beta^{\mathrm{mn}}(\mathbf{k},\mathbf{q}) &= \sum_{\alpha\alpha'} C_{\mathrm{m}\mathbf{k}}^{\alpha*} C_{\mathrm{n},\mathbf{k}+\mathbf{q}}^{\alpha'} \sum_{\mathbf{R},\mathbf{R}'} e^{-\mathbf{k}(\mathbf{R}+\boldsymbol{\tau}_\alpha)} e^{-(\mathbf{k}+\mathbf{q})(\mathbf{R}'+\boldsymbol{\tau}'_\alpha)} \\ &\quad \times \int \phi_\alpha(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_\alpha)\phi_\beta(\mathbf{r}-\boldsymbol{\tau}_\beta)\phi_{\alpha'}(\mathbf{r}-\mathbf{R}'-\boldsymbol{\tau}'_\alpha)d^3r \quad . \end{aligned} \tag{3.14}$$

Die auftretenden Dreizentrenintegrale $\int \phi_\alpha\phi_\beta\phi_{\alpha'}$ von drei Gauß-Orbitalen an drei verschiedenen Aufpunkten lassen sich rekursiv aus dem Integral $\int \phi_{\alpha,\mathrm{s}}\phi_{\beta,\mathrm{s}}\phi_{\alpha',\mathrm{s}}$ der drei s-Orbitale bestimmen. Diese äußerst effiziente Methode, die von Obara und Saika [Oba86] vorgeschlagen wurde, stellen wir in Anhang C dar.

3.3 Dielektrische Matrix

Ausgehend von der nach Gl. (3.12) berechneten Polarisierbarkeit P, erhält man durch eine Faltung mit der Coulomb-Wechselwirkung v die dielektrische Matrix ϵ:

$$\epsilon = 1 - vP \quad . \tag{3.15}$$

Bei dieser dielektrischen Funktion treten Matrixelemente auf, die für $\mathbf{q}\to 0$ wie $1/|\mathbf{q}|$ divergieren. Diese Divergenz umgehen wir, indem wir mit einer *symmetrisierten dielektrischen Matrix*

$$\tilde{\epsilon} = 1 - \tilde{v}P\tilde{v} \tag{3.16}$$

arbeiten. Die Hilfsfunktion $\tilde{v}(\mathbf{r},\mathbf{r}') = e\pi^{-3/2}/|\mathbf{r}-\mathbf{r}'|^2$ ersetzt hierbei die Coulomb-Wechselwirkung. Der Zusammenhang zwischen der ursprünglichen und der symmetrisierten dielektrischen Matrix sowie die Eigenschaften der Funktion $\tilde{v}$ werden in Anhang D ausführlich dargestellt.

Im von uns verwendeten Funktionenraum $\{\chi_\beta\}$ wird der in Gl. (3.16) skizzierte Rechenschritt folgendermaßen realisiert:

$$[\tilde{\epsilon}](\mathbf{q}) = S(\mathbf{q}) - [\tilde{v}](\mathbf{q}) \cdot S^{-1}(\mathbf{q}) \cdot [P](\mathbf{q}) \cdot S^{-1}(\mathbf{q}) \cdot [\tilde{v}](\mathbf{q}) \quad . \tag{3.17}$$

In dieser Gleichung treten als Folge der Symmetrisierung keine divergenten Matrixelemente mehr auf. Jedoch führt der langreichweitige Charakter der Coulomb-Wechselwirkung zu zwei anderen numerischen Schwierigkeiten, die eine sehr sorgfältige Behandlung erfordern. Die eine Schwierigkeit besteht in der Bestimmung der Matrixelemente $[\tilde{v}]_{\beta\beta'}(\mathbf{q})$ der Hilfsfunktion. Dies wird in Anhang D erläutert. Das andere Problem besteht darin, die dielektrische Matrix $[\tilde{\epsilon}](\mathbf{q})$ für $\mathbf{q}\to\mathbf{0}$ stabil gegen ihren richtigen Grenzwert konvergieren zu lassen, was zur Berechnung der dielektrischen Konstanten erforderlich ist. Hierzu verweisen wir auf Anhang E.

Die mit diesen Methoden berechnete dielektrische Matrix wird invertiert (vgl. Anhang A):

$$\langle\tilde{\epsilon}^{-1}\rangle = [\tilde{\epsilon}]^{-1} \quad . \tag{3.18}$$

Die inverse dielektrische Matrix erlaubt die Bestimmung der makroskopischen dielektrischen Funktion ϵ_M. Ausgehend von der Fourierreihe für $\tilde{\epsilon}^{-1}$, ist diese Größe durch den Kopf der inversen dielektrischen Matrix gegeben:

$$\epsilon_M(\mathbf{q},\omega) = \left(\tilde{\epsilon}^{-1}_{\mathbf{00}}(\mathbf{q},\omega)\right)^{-1} \quad . \tag{3.19}$$

Wenn die inverse dielektrische Matrix in der Gauß-Basis gegeben ist, so ergibt sich

$$\epsilon_M(\mathbf{q},\omega) = \left(\sum_{\beta\beta'} K^{\mathbf{0}}_\beta(\mathbf{q})\langle\tilde{\epsilon}^{-1}\rangle_{\beta,\beta'}(\mathbf{q},\omega)\left(K^{\mathbf{0}}_{\beta'}(\mathbf{q})\right)^*\right)^{-1} \quad . \tag{3.20}$$

[$K^{\mathbf{G}}_\beta(\mathbf{q})$ ist die Fouriertransformierte des Gauß-Orbitals β (vgl. Anhang A).]

Eine erneute Faltung der inversen dielektrischen Matrix mit der Hilfsfunktion $\tilde{v}$ führt auf die abgeschirmte Wechselwirkung W:

$$[W] = [\tilde{v}\tilde{\epsilon}^{-1}\tilde{v}] \tag{3.21}$$

$$= [\tilde{v}]S^{-1}[\tilde{\epsilon}^{-1}]S^{-1}[\tilde{v}] \tag{3.22}$$

$$= [\tilde{v}]\langle\tilde{\epsilon}^{-1}\rangle[\tilde{v}] \quad . \tag{3.23}$$

Nun sind wir in der Lage, die gewünschten Matrixelemente des Selbstenergie-Operators für eine noch einzusetzende Energie E (meistens $E{=}E_{\mathrm{m'k}}$) zu bestimmen:

$$\langle\psi_{\mathrm{mk}}|\Sigma(E)|\psi_{\mathrm{m'k}}\rangle = \sum_{\mathbf{q}}\sum_{\mathrm{n}}\sum_{\beta\beta'} I_{\beta}^{\mathrm{mn}}(\mathbf{k},\mathbf{q})\left(I_{\beta'}^{\mathrm{m'n}}(\mathbf{k},\mathbf{q})\right)^{*}$$

$$\times\frac{i}{2\pi}\int e^{i\omega 0^{+}}\frac{\langle\tilde{\epsilon}^{-1}_{\beta\beta'}\rangle(-\mathbf{q},\omega)}{E-\omega-E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}+i0^{+}\mathrm{sign}(E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}-E_F)}d\omega \tag{3.24}$$

$$\text{mit} \quad I_{\beta}^{\mathrm{mn}}(\mathbf{k},\mathbf{q}) = \sum_{\beta'\beta''} M_{\beta'}^{\mathrm{mn}}(\mathbf{k},\mathbf{q})S^{-1}_{\beta'\beta''}(-\mathbf{q})[\tilde{v}]_{\beta''\beta}(-\mathbf{q}) \quad . \tag{3.25}$$

Die verbleibende Frequenzintegration $\int d\omega$ erweist sich bei näherer Betrachtung als schwierig zu behandeln. Eine direkte numerische Bestimmung stellt aus folgenden Gründen einen erheblichen Aufwand dar. Die Funktion $1/(\Delta E-\omega\pm i0^{+})$ besitzt einfache Polstellen auf der reellen Frequenzachse. Eine konvergente numerische Integration verlangt also eine Vielzahl an Frequenz-Stützstellen. Es erfordert einen enormen Rechenaufwand, die dielektrische Matrix für die notwendige große Zahl von Frequenzen auszuwerten. Weiterhin stellt sich heraus, daß diese dielektrische Matrix auf der reellen Frequenzachse selber eine polartige Struktur aufweist, die durch Plasmonen gekennzeichnet ist. Diese Struktur nehmen wir zum Anlaß, die dielektrische Matrix durch ein entsprechendes *Plasmon-Pol-Modell* zu modellieren. Dadurch läßt sich die Frequenzintegration in einfacher Weise analytisch ausführen. Dies wird in Kap. 4 diskutiert.

Es ist zu beachten, daß die in Gl. (3.9) und (3.24) verwendete Green-Funktion, die einem System effektiv nicht wechselwirkender Teilchen entspricht, im Grunde ebenfalls als *„Elektronen-Pol-Modell“* verstanden werden muß. Die tatsächliche Energieabhängigkeit der Green-Funktion ist wesentlich komplexer. Wir werden diesen Aspekt in Kap. 14 diskutieren.

3.4 Ergebnisse für Si

Bevor wir uns dem Plasmon-Pol-Modell zuwenden, werden wir zunächst in diesem Abschnitt am Beispiel des Si Volumenkristalls nachweisen, daß die Verwendung von Gauß-Orbitalen als GW-Basissatz zu den gleichen Ergebnissen führt wie Rechnungen mit ebenen Wellen. Gleichzeitig stellen wir die wesentlichen Eigenschaften von Quasiteilchen-Korrekturen zu LDA-Bandstrukturen vor.

Wir führen zunächst eine LDA-Rechnung für Si durch. Dabei kommen normerhaltende ab-initio Pseudopotentiale (PP) zur Anwendung. Wir verwenden in dieser Arbeit sowohl nichtseparable [Ham79, Bac82, Ham91] als auch separable [Kle82, Stu90] PP. Das Austausch-Korrelationsfunktional beruht auf den Monte-Carlo Daten von Ceperley und Alder [Cep80] für das homogene Elektronengas in der von Perdew und Zunger [Per81] parametrisierten Form. Wir stellen die Wellenfunktionen mit Hilfe von Gauß-Orbitalen dar (GO-LDA; Ablingkonstanten: 0.15 und 0.5). Die theoretische Gitterkonstante ergibt sich mit diesen Rechnungen zu 5.37 Å. Sie ist etwas kleiner als der experimentell ermittelte Wert von 5.43 Å. Diese Unterschätzung der Gitterkonstanten um ca. 1 % ist typisch für LDA-Rechnungen unter Verwendung des Austausch-Korrelations-Potentials von Ceperley und Alder. [Andere Ausdrücke für das Austausch-Korrelations-Funktional, etwa der von Wigner [Wig37], mögen zwar auf Gitterkonstanten führen, die mit den experimentellen Daten besser übereinstimmen. Die leichte Unterschätzung der Gitterkonstanten ist jedoch eine Eigenschaft der LDA selbst. Das Funktional von Ceperley und Alder kommt der wirklichen LDA sicherlich am nächsten und wird deshalb von uns durchgehend verwendet.] Wir werden in Teil II der Arbeit Oberflächen von Si untersuchen. Als Oberflächengeometrien werden wir die Ergebnisse von LDA-Strukturoptimierungen zugrunde legen, die grundsätzlich auf der Basis der theoretischen Gitterkonstante durchzuführen sind. Daher führen wir schon hier alle Rechnungen für Si bei der theoretischen Gitterkonstante aus. Die fundamentale Energielücke in der Bandstruktur beträgt 0.46 eV und ist, wie aufgrund des allgemeinen Gap-Problems der DFT zu erwarten ist, kleiner als der experimentelle Wert von 1.17 eV. Wir merken an dieser Stelle an, daß die Energielücke von der Gitterkonstante abhängt. Mit kleiner werdender Gitterkonstante nimmt das Gap um etwa 0.1 eV je 0.1 Å ab. [Dies entspricht der experimentellen Beobachtung, daß die Energielücke von Si unter Druck kleiner wird.] Die bei der theoretischen Gitterkonstante berechnete Gap-Energie ist um 0.06 eV kleiner als der bei der experimentellen Gitterkonstante bestimmte Wert.

Aufbauend auf dieser LDA-Rechnung, führen wir zwei separate GW-Rechnungen durch. In der einen verwenden wir einen herkömmlichen Basissatz aus 137 ebenen Wellen (PW-GWA) [Roh93a, Roh93b]. Für die andere Rechnung benutzen wir hingegen Gauß-Orbitale als GW-Basisfunktionen (GO-GWA). Wir verwenden die Abklingkonstanten 0.15 und 0.5. Zu jeder Abklingkonstanten zählen zehn Orbitale (s-, p-, d- und s^*-Typ). Es stellt sich heraus, daß darüber hinaus noch f-Orbitale notwendig sind, um eine ausreichende Flexibilität des Basissatzes zu gewährleisten. [Ohne f-Orbitale ergeben sich Fehler von bis zu 0.5 eV in der Quasiteilchen-Bandstruktur.] Wir erzielen diese Flexibilität, indem wir entsprechende kartesische Gauß-Funktionen der Klasse L=3 mit der Abklingkonstante 0.5 zum Basissatz hinzunehmen (siehe Anhang B). An jedem Atom sind nunmehr 30 Basisfunktionen zentriert: zehn Funktionen (s, p, d, s^*) mit der Abklingkonstante 0.15 und zwanzig Funktionen (s, p, d, s^* und f) mit der Abklingkonstante 0.5. Da die Einheitszelle zwei Atome enthält, ergibt sich letztlich ein GW-Basissatz aus 60 Orbitalen.

Eine konvergente LDA-Rechnung für Si im Rahmen der PW-LDA erfordert etwa 400 ebene Wellen als Basissatz (entsprechend einer Abschneideenergie von 15–20 Ryd). Eine GW-Rechnung im Rahmen der PW-GWA erfordert eine geringere Zahl von ebenen Wellen (137 ebene Wellen für eine konvergente Behandlung). Die Ursache liegt darin, daß der Selbstenergie-Operator räumlich schwächer strukturiert ist als die Potentiale der LDA. Bei den Gauß-Orbital-Basen ist hingegen zu beobachten, daß für die GO-GWA *mehr* Ansatzfunktionen notwendig sind (für Si: 60 GO) als für die GO-LDA (40 GO), um eine konvergente Bandstruktur zu erzielen. Das liegt daran, daß der Selbstenergie-Operator und die in ihn eingehenden Größen eine kompliziertere Drehimpulsstruktur aufweisen als die Wellenfunktionen. [Daraus ergibt sich die Notwendigkeit, höhere Drehimpulskomponenten als Ansatzfunktionen zu berücksichtigen.] Dennoch ist die Anzahl an GO-Basisfunktionen für die GO-GWA erheblich geringer als die Anzahl an ebenen Wellen für die PW-GWA, um die gleiche Basissatz-Konvergenz zu erzielen.

Von den unterschiedlichen Basissätzen abgesehen, führen wir die beiden Rechnungen in der gleichen Weise durch. Die volle statische dielektrische Matrix wird in der RPA bestimmt und mit einem Plasmon-Pol-Modell zur dynamischen Matrix erweitert. [Wie dieses Plasmon-Pol-Modell in der GO-GWA zu konstruieren ist, erläutern wir in Kap. 4.] Zur Bestimmung der Zweipunkt-Funktionen verwenden wir zwei spezielle Punkte [Cha73]. Zur Berechnung der statischen dielektrischen Matrix für einen gegebenen $\mathbf{q}$-Punkt verwenden wir ebenfalls zwei spezielle Punkte. Ledig-

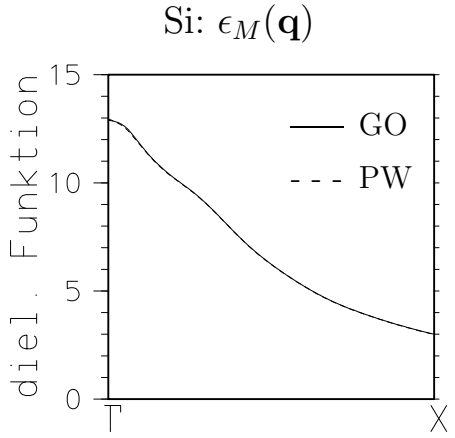


Abbildung 3.1: *Statische makroskopische dielektrische Funktion $\epsilon_M(\mathbf{q},\omega{=}0)$ von Si entlang der Hochsymmetrielinie Γ—X. Die durchgezogene Kurve wurde mit 60 Gauß-Orbitalen als Basissatz für die dielektrische Matrix berechnet, die gestrichelte Kurve mit 137 ebenen Wellen. Die beiden Kurven liegen fast aufeinander.*

lich die Bestimmung des Kopfes der dielektrischen Matrix erfordert im Limes $\mathbf{q}{\to}\mathbf{0}$ zehn spezielle Punkte.

Im folgenden vergleichen wir für verschiedene physikalische Größen, die im Rahmen der Rechnungen auftreten, die Ergebnisse der GO-GWA mit denen der PW-GWA. In Abb. 3.1 stellen wir die statische makroskopische dielektrische Funktion $\epsilon_M(\mathbf{q},\omega{=}0)$ von Si dar. Die beiden Kurven wurden in völlig gleicher Weise bestimmt, und zwar zum einen mit dem erwähnten Basissatz aus Gauß-Orbitalen (durchgezogene Kurve), zum anderen mit ebenen Wellen (gestrichelte Kurve). Offensichtlich ergeben die beiden Basissätze genau die gleiche makroskopische dielektrische Funktion; diesbezüglich sind beide Basissätze gleichwertig. In Abb. 3.2 stellen wir das statische Abschirmpotential $W(\mathbf{r},\mathbf{r}')$–$v(\mathbf{r},\mathbf{r}')$ dar, das bei $\mathbf{r}$ durch ein zusätzliches Elektron, das sich an der Stelle $\mathbf{r}'$ befindet, induziert wird. Auch für diese Funktion stimmen die Ergebnisse des PW-Basissatzes (gestrichelte Linien) hervorragend mit denen der GO-Basis (durchgezogene Linien) überein.

Als letzten Schritt bestimmen wir die Quasiteilchenkorrekturen (2.19) zu den LDA-Einteilchenenergien der Zustände in Si. Bevor wir auch hier die PW-GWA mit den GO-GWA-Resultaten vergleichen, stellen wir die Eigenschaften der Korrekturen und ihre Auswirkungen auf die Bandstruktur vor. Die in der GO-GWA-Rechnung bestimmten Korrekturen sind in Abb. 3.3 in Abhängigkeit von der jeweiligen LDA-Bandstruktur-Energie dargestellt. Die Korrekturen sind für Valenzband-

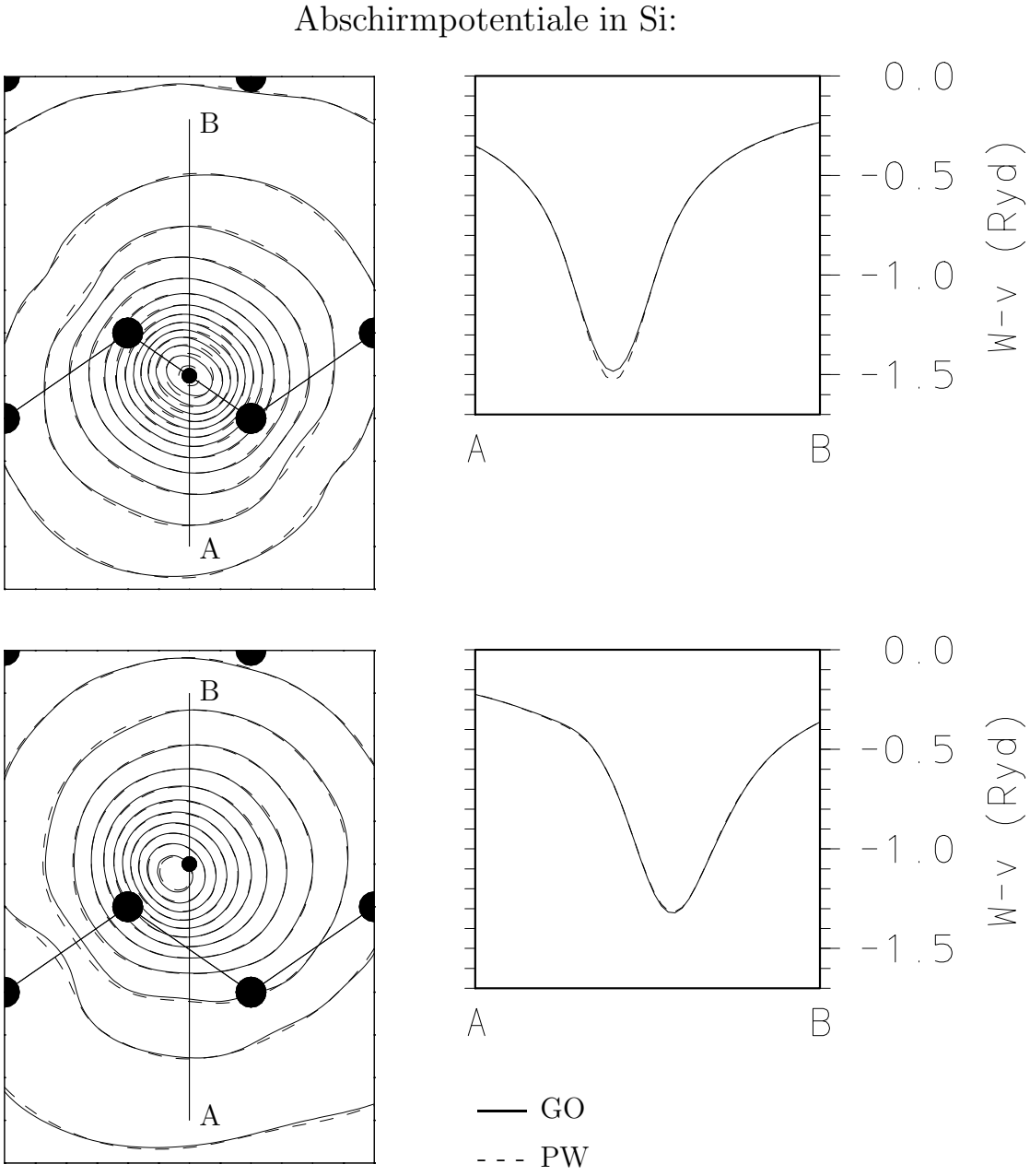


Abbildung 3.2: *Abschirmpotential $W - v$, das durch ein Elektron (kleiner Punkt) in der $[1\bar{1}0]$-Ebene in Si induziert wird. Die gestrichelten Linien stellen das Ergebnis einer Ebene-Wellen-Rechnung dar, die durchgezogenen Linien eine Gauß-Orbital-Rechnung. Der Niveaulinienabstand beträgt 0.1 Ryd. Auf der rechten Seite werden die Abschirmpotentiale entlang der Linie $\overline{AB}$ dargestellt.*

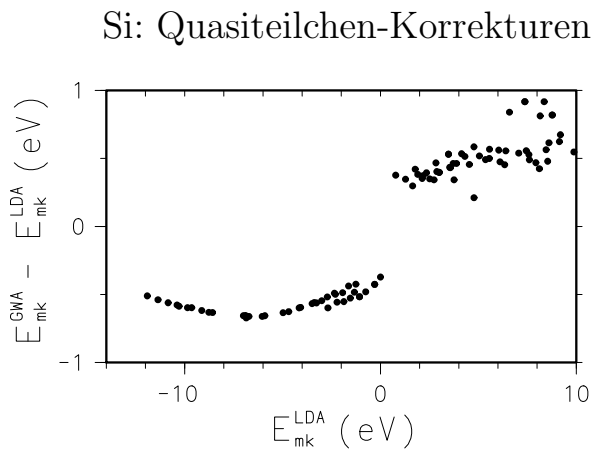


Abbildung 3.3: *Quasiteilchen-Korrekturen* $E^{\mathrm{GWA}}_{\mathbf{mk}} - E^{\mathrm{LDA}}_{\mathbf{mk}} = \langle \mathrm{m}\mathbf{k}|\Sigma(E^{\mathrm{GWA}}_{\mathbf{mk}}) - V_{\mathrm{xc}}|\mathrm{m}\mathbf{k}\rangle$ *für Si. Die Korrekturen sind gegen die LDA-Energie des jeweiligen Zustandes aufgetragen.*

zustände ($E^{\mathrm{DFT}}_{\mathbf{mk}}<0$) negativ und für Leitungsbandzustände ($E^{\mathrm{DFT}}_{\mathbf{mk}}>0$) positiv. Die Korrektur der Valenzbänder ist für alle Zustände nahezu konstant; sie reicht von –0.65 eV bis –0.4 eV. Die Korrekturdaten der Leitungsbänder streuen etwas mehr; sie liegen im Bereich von +0.2 eV bis +1.0 eV. Das Valenzbandmaximum erfährt eine Verschiebung von –0.41 eV nach unten, während das Leitungsbandminimum um +0.36 eV angehoben wird. Dadurch erhöht sich das fundamentale Gap um +0.77 eV gegenüber der LDA und beträgt nunmehr 1.23 eV. Dieser Wert stimmt mit dem experimentellen Resultat von 1.17 eV hervorragend überein. Das Gap-Problem der DFT wird durch diesen Zugang überwunden.

Im linken Teil der Abb. 3.4 stellen wir die entsprechenden Bandstrukturen dar, wie sie sich aus der LDA (gepunktete Kurven) und aus der GO-GWA (durchgezogene Kurven) ergeben. [Es ist zu beachten, daß für beide Bandstrukturen das jeweilige Valenzbandmaximum als 0 eV festgelegt wurde.] Es wird deutlich, daß durch die Korrekturen die Dispersion der Valenzbänder einerseits und die der Leitungsbänder andererseits nahezu unverändert bleiben. Als wesentlicher Effekt der Korrekturen wird der gesamte Leitungsbandbereich fast starr zu höheren Energien verschoben. Im weiteren Verlauf der Arbeit werden auch Systeme auftreten, bei denen die Dispersion einzelner Bänder oder die Lage verschiedener Valenzbänder zueinander durch die Quasiteilchen-Korrekturen deutlich verändert werden. Die hervorragende Übereinstimmung der GW-Bandstruktur mit den Ergebnissen von ARPES- und

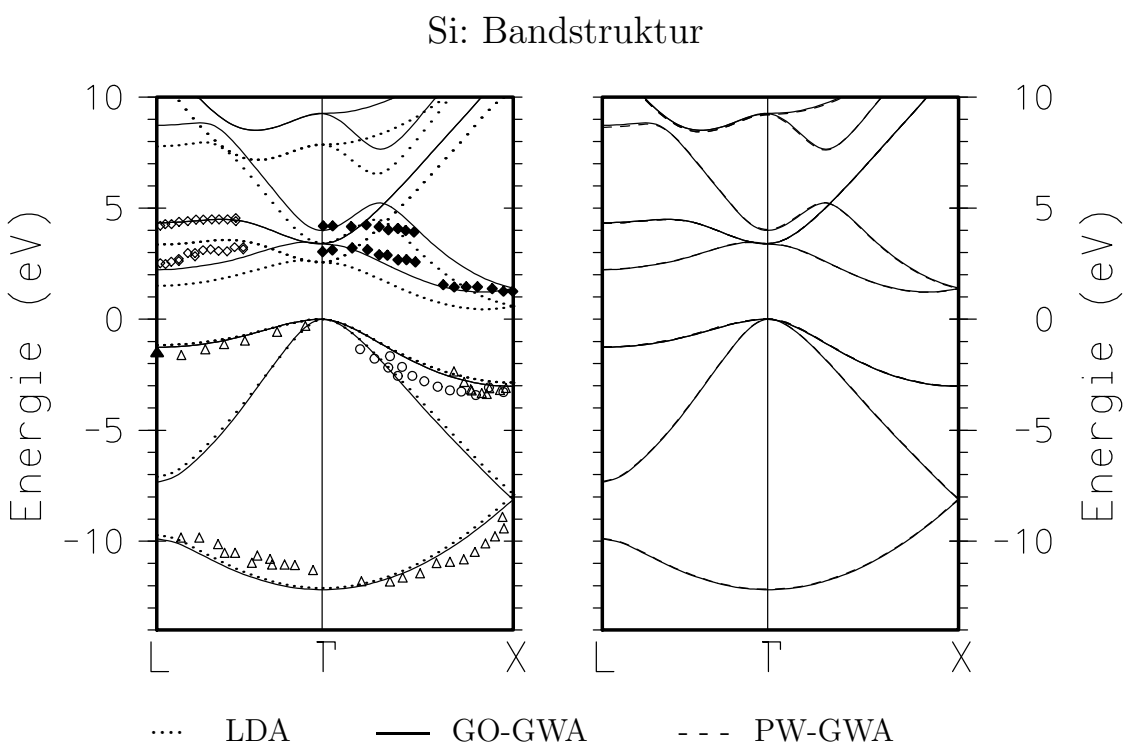


Abbildung 3.4: *Linke Seite: Vergleich der GW-Bandstruktur für Si mit der zugrunde liegenden LDA-Bandstruktur (gepunktet) und mit experimentellen ARPES- und ARIPES-Daten (einzelne Symbole; vgl. [Ort93] und dortige Referenzen). Rechte Seite: Vergleich der GO-GWA-Bandstruktur für Si (durchgezogene Kurven) mit der PW-GWA-Bandstruktur (gestrichelt).*

ARIPES-Experimenten ([Ort93] und dortige Referenzen; einzelne Symbole in Abb. 3.4) bestätigt die Leistungsfähigkeit der GWA bei der Bestimmung von Einteilchen-Energien. Im rechten Teil der Abbildung stellen wir die Quasiteilchen-Bandstrukturen der PW-GWA Rechnung (gestrichelte Kurven) und der GO-GWA Rechnung (durchgezogene Kurven) einander gegenüber. Die beiden Bandstrukturen unterscheiden sich nur marginal. Die größten Abweichungen finden sich in den höheren Leitungsbändern; sie betragen maximal 0.08 eV. Auch die absoluten Werte der Quasiteilchen-Korrekturen stimmen hervorragend miteinander überein; im rechten Teil der Abb. 3.4 ist keine Verschiebung der Valenzbandmaxima gegeneinander enthalten.

Zusammenfassend stellen wir fest, daß die beiden Basissätze (Gauß-Orbitale und ebene Wellen) zur Darstellung der Zweipunktfunktionen der GW-Näherung und zur Berechnung von Quasiteilchen-Korrekturen beide gleichermaßen gut geeignet

sind. Beide Basissätze führen zu exakt gleichen Ergebnissen. Das ausschlaggebende Kriterium für die Verwendung des einen oder des anderen Basissatzes bei der Behandlung eines konkreten Systems wird nunmehr der numerische Aufwand sein. Generell wird sich bei Systemen mit schwach veränderlicher Ladungsdichte die Verwendung von ebenen Wellen als günstiger erweisen. Die herkömmliche Behandlung von Volumen-Halbleitern im Rahmen eines konventionellen Pseudopotentialkonzeptes, wie etwa die hier vorgestellte Rechnung für Si, läßt sich mit ebenen Wellen und mit Gauß-Orbitalen gleichermaßen gut durchführen. Bei Oberflächensystemen (insbesondere bei solchen mit orthogonalen Gittervektoren; vgl. Abschnitt 7.2) und bei Systemen mit stark inhomogener Ladungsdichte ist hingegen die Verwendung von Gauß-Orbitalen vorteilhaft; in vielen Fällen wird hierdurch die Durchführung einer GW-Studie überhaupt erst möglich.

Kapitel 4

Das Plasmon-Pol-Modell

Wir wenden uns nun dem bereits erwähnten Plasmon-Pol Modell zur Beschreibung der zeitlichen Dynamik der Abschirmung zu. Das Modell ermöglicht die Auswertung der Frequenzintegration in der Selbstenergie (3.24). Eine direkte Auswertung dieser Frequenzintegration ist mit erheblichem Aufwand verbunden. Es bestehen im wesentlichen zwei Möglichkeiten, die Frequenzintegration in Gl. (3.24) zu vereinfachen.

Zum einen läßt sich für Energien E in der Nähe des fundamentalen Gaps der Integrationsweg $\mathcal{C}$ auf einen Weg $\mathcal{C}'$ entlang der imaginären Frequenzachse verlagern (siehe linke Figur in Abb. 4.1). Der neue Integrationsweg vermeidet die Pole der Funktionen G und W vollständig. Beide Funktionen sind entlang $\mathcal{C}'$ so schwach strukturiert, daß eine numerische Integration mit wenigen Stützstellen möglich wird. Godby *et al.* haben auf diese Weise mit lediglich zehn Stützstellen die entsprechenden Frequenzintegrale berechnet [God88]. Die Gültigkeit der Transformation $\mathcal{C} \to \mathcal{C}'$ ist für reelle Energien E auf das Intervall $E_{\mathrm{VBM}} - E_g < E < E_{\mathrm{CBM}} + E_g$ beschränkt. Bei anderen Energien überlappen die Polstellen der Green-Funktion und die der abgeschirmten Wechselwirkung (siehe rechte Figur in Abb. 4.1), so daß die Verlagerung $\mathcal{C} \to \mathcal{C}'$ nicht mehr exakt möglich ist. Eng verwandt mit der Transformation $\mathcal{C} \to \mathcal{C}'$ ist die Vorgehensweise, die Selbstenergie von vornherein für *komplexe* Energien $E = E_{\mathrm{VBM}} + 1/2E_g + i\varepsilon$ durch Integration entlang $\mathcal{C}'$ zu berechnen und sie dann als analytische Fortsetzung auf die reelle Energieachse zu transformieren [Roj95].

Die andere Möglichkeit besteht in der Verwendung von Plasmon-Pol-Modellen. Diesen Modellen wenden wir uns nun zu.

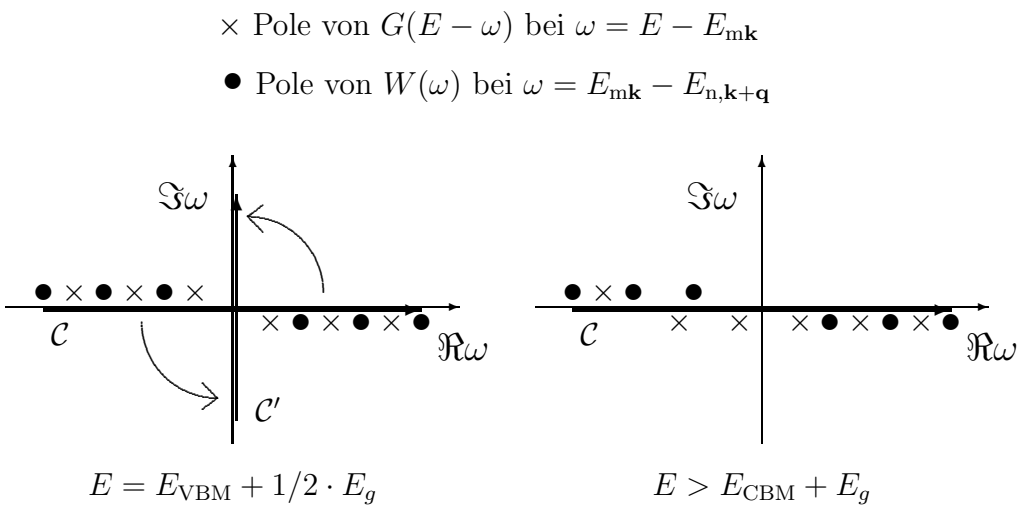


Abbildung 4.1: *Die Verlagerung des Frequenzintegrals zur Selbstenergie vom Weg $\mathcal{C}$ (entlang der reellen Achse) auf den Weg $\mathcal{C}'$ (in der komplexen Ebenen) ist nur für Energien E mit $E_{\mathrm{VBM}} - E_g < E < E_{\mathrm{CBM}} + E_g$ exakt (linke Seite). Für Energien E außerhalb dieses Intervalls überdecken sich die Pole von G und W auf der reellen Achse, so daß eine Transformation $\mathcal{C} \to \mathcal{C}'$ nicht mehr residuenfrei möglich ist (rechte Seite).*

4.1 Einfache Plasmon-Pol Modelle

Wir verfolgen in dieser Arbeit den zweiten Zugang zur Auswertung der Frequenzintegration in (3.24). Wir gehen davon aus, daß die Frequenzabhängigkeit der inversen dielektrischen Matrix im wesentlichen durch Pole der Form

$$\frac{1}{\omega^2 - \bar{\omega}^2} \tag{4.1}$$

beschrieben werden kann. Das einfachste Beispiel hierfür bietet die inverse dielektrische Funktion des homogenen Elektronengases im Rahmen der Random-Phase Approximation (RPA):

$$\epsilon^{-1}(q,\omega) = 1 + \frac{\bar{\omega}^2(q)}{\omega^2 - \bar{\omega}^2(q)} \quad . \tag{4.2}$$

Bei der Plasmonfrequenz

$$\bar{\omega}(q) = \omega_{\mathrm{Pl}} \left(1 + \frac{3k_F^2}{5m^2\omega_{\mathrm{Pl}}^2} q^2 + \ldots \right)^{1/2} \tag{4.3}$$

divergiert die inverse dielektrische Funktion. Das ermöglicht das Auftreten von *Plasmonen*, also von Eigenschwingungen des Elektronensystems als kollektive Anregung in Form einer ebenen Welle mit dem Wellenvektor q.

Es stellt sich heraus, daß die für den Selbstenergie-Operator wesentlichen Abschirm-Eigenschaften durch diese Plasmonen erfaßt werden. Darauf aufbauend, stellen wir in dieser Arbeit die dielektrische Matrix in Form von Plasmonen dar. Dieses sogenannte *Plasmon-Pol-Modell* ermöglicht die analytische Durchführung der Frequenzintegration in Gl. (3.24). Plasmon-Pol-Modelle in Ebene-Wellen-Darstellung wurden mit großem Erfolg zur Berechnung der Quasiteilchen-Bandstruktur für eine Vielzahl von Materialien eingesetzt [Hyb86, Hot88, Zhu91, Roh93a, Roh93b]. Es gilt nun, dieses Konzept auf die von uns verwendete Gauß-Basis zu übertragen. Das einfachste in dieser Arbeit benutzte Plasmon-Pol-Modell besitzt die Form

$$\langle \tilde{\epsilon}^{-1} \rangle_{\beta\beta'}(\mathbf{q},\omega) \;=\; \sum_l \xi^l_\beta(\mathbf{q}) \lambda^{-1}_{\mathbf{q}l}(\omega) \left(\xi^l_{\beta'}(\mathbf{q})\right)^* \tag{4.4}$$

$$\text{mit} \quad \lambda^{-1}_{\mathbf{q}l}(\omega) \;=\; 1 + \frac{z_{\mathbf{q}l}\omega_{\mathbf{q}l}}{2} \left(\frac{1}{\omega - (\omega_{\mathbf{q}l} - i0^+)} - \frac{1}{\omega + (\omega_{\mathbf{q}l} - i0^+)} \right) \quad . \tag{4.5}$$

Dabei stellen die Größen ξ^l_β und $\lambda^{-1}_{\mathbf{q}l}(\omega)$ verallgemeinerte Eigenvektoren und -werte der inversen dielektrischen Matrix dar (siehe unten). In diesem Ansatz gehen wir davon aus, daß nur die Eigen*werte* $\lambda^{-1}_{\mathbf{q}l}(\omega)$ von der Frequenz ω abhängen, während wir die Eigen*vektoren* ξ^l_β als frequenzunabhängig annehmen. [Das stellt natürlich eine gewisse Einschränkung für die inverse dielektrische Matrix dar]. Die inverse dielektrische Matrix wird für alle Frequenzen durch die Größen $\xi^l_\beta(\mathbf{q})$, $z_{\mathbf{q}l}$ und $\omega_{\mathbf{q}l}$ festgelegt. $\xi^l_\beta(\mathbf{q})$ und $z_{\mathbf{q}l}$ gewinnen wir aus der *statischen* dielektrischen Matrix, d.h. für ω=0. Analog zum Plasmon-Pol-Modell in PW-Notation [Lin88, Roh93a, Roh93b] führen wir auch hier eine Diagonalisierung einer geeigneten Matrix durch. Der Einfachheit halber wählen wir zur Darstellung der Vorgehensweise eine Matrix-Notation und verzichten auf das Argument $\mathbf{q}$. Als Hilfsgröße benötigen wir die Choleski-Zerlegung der Überlapp-Matrix $S = LL^\dagger$. Damit konstruieren wir aus der nach (3.17) berechneten dielektrischen Matrix $[\tilde{\epsilon}]$ eine Matrix, die die gleichen Eigenwerte λ aufweist wie die entsprechende dielektrische Matrix in Ebene-Wellen-Notation, und die sich diagonalisieren läßt (Eigenwerte λ_l, Eigenvektoren ϕ^l_β):

$$L^{-1}\ [\tilde{\epsilon}](\omega{=}0)\ L^{-1^\dagger} \equiv \Phi\ \Lambda(\omega{=}0)\ \Phi^\dagger \tag{4.6}$$

$$\text{mit} \quad \Phi = \left\{\phi^l_\beta\right\} \quad \text{und} \quad \Lambda = \{\lambda_l \delta_{ll'}\} \quad .$$

Zwischen den Eigenwerten λ_l und den Parametern z_l gilt der gleiche Zusammenhang wie in PW-Notation [Roh93a, Roh93b]:

$$z_l = 1 - \frac{1}{\lambda_l(\omega = 0)} \quad . \tag{4.7}$$

Nun läßt sich die inverse dielektrische Matrix für *alle* Frequenzen konstruieren:

$$\begin{aligned}\langle\tilde{\epsilon}^{-1}\rangle(\omega) &= [\tilde{\epsilon}]^{-1}(\omega) = \left(L\Phi\Lambda(\omega)\Phi^\dagger L^\dagger\right)^{-1} = L^{-1^\dagger}\Phi\Lambda^{-1}(\omega)\Phi^\dagger L^{-1} \\ &=: \Xi\Lambda^{-1}(\omega)\Xi^\dagger \quad . \end{aligned} \tag{4.8}$$

[Man beachte, daß die Eigenvektoren eine orthogonale Matrix Φ bilden.] Die in Gl. (4.8) auftretenden Vektoren ξ^l_β berechnen sich gemäß

$$\xi^l_\beta = \sum_{\beta'}(L^{-1}_{\beta'\beta})^*\phi^l_{\beta'} \quad . \tag{4.9}$$

Zur Festlegung der noch unbestimmten Parameter $\omega_{\mathbf{q}l}$ ziehen wir in der Regel die f-Summenregel [Joh74] heran:

$$\frac{1}{2\pi i}\int\limits_{-\infty}^{\infty}\omega\left[\tilde{\epsilon}^{-1}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega)-\delta_{\mathbf{G},\mathbf{G}'}\right]d\omega = -\frac{\omega^2_{\mathrm{Pl}}}{2}\frac{(\mathbf{q}+\mathbf{G})}{|\mathbf{q}+\mathbf{G}|}\frac{(\mathbf{q}+\mathbf{G}')}{|\mathbf{q}+\mathbf{G}'|}\frac{\rho(\mathbf{G}-\mathbf{G}')}{\rho(\mathbf{0})} \quad . \tag{4.10}$$

Setzt man für $\tilde{\epsilon}^{-1}$ das Plasmon-Pol-Modell ein, so ergeben sich die Parameter $\omega_{\mathbf{q}l}$ zu

$$\omega^2_{\mathbf{q}\mathrm{l}} = \frac{\omega^2_{\mathrm{Pl}}}{z_{\mathbf{q}\mathrm{l}}}\sum_{\mathbf{G},\mathbf{G}'}\left(\sum_\beta\xi^l_\beta(\mathbf{q})K^{\mathbf{G}}_\beta(\mathbf{q})\right)^*\frac{(\mathbf{q}+\mathbf{G})}{|\mathbf{q}+\mathbf{G}|}\frac{(\mathbf{q}+\mathbf{G}')}{|\mathbf{q}+\mathbf{G}'|}\frac{\rho(\mathbf{G}-\mathbf{G}')}{\rho(\mathbf{0})}\left(\sum_\beta\xi^l_\beta(\mathbf{q})K^{\mathbf{G}'}_\beta(\mathbf{q})\right) \tag{4.11}$$

[$K^{\mathbf{G}}_\beta(\mathbf{q})$ ist die Fouriertransformierte des Gauß-Orbitals β (vgl. Anhang A).] Das einfache analytische Frequenzverhalten (4.5) der inversen dielektrischen Matrix im Rahmen des Plasmon-Pol-Modells ermöglicht die analytische Auswertung des Frequenzintegrals in (3.24). Die Matrixelemente der Selbstenergie lauten somit

$$\begin{aligned}\langle\psi_{\mathrm{m}\mathbf{k}}|\Sigma(E)|\psi_{\mathrm{m}'\mathbf{k}}\rangle &= \sum_{\mathbf{q}}\sum_{\mathrm{n}}\sum_l J^{\mathrm{mn}}_l(\mathbf{k},\mathbf{q})\left[J^{\mathrm{m'n}}_l(\mathbf{k},\mathbf{q})\right]^* \\ &\quad\times\begin{cases}\left(-1+\dfrac{z_{-\mathbf{q}\mathrm{l}}\omega_{-\mathbf{q}\mathrm{l}}}{2}\dfrac{1}{E-E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}+\omega_{-\mathbf{q}\mathrm{l}}}\right) & \text{für n}\in\text{Val} \\ \left(\dfrac{z_{-\mathbf{q}\mathrm{l}}\omega_{-\mathbf{q}\mathrm{l}}}{2}\dfrac{1}{E-E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}-\omega_{-\mathbf{q}\mathrm{l}}}\right) & \text{für n}\in\text{Con} \quad .\end{cases}\end{aligned} \tag{4.12}$$

$$\text{mit} \quad J^{\mathrm{mn}}_l(\mathbf{k},\mathbf{q}) = \sum_\beta I^{\mathrm{mn}}_\beta(\mathbf{k},\mathbf{q})\xi^l_\beta(-\mathbf{q}) \quad . \tag{4.13}$$

Die in Gl. 4.12 auftretende Brillouinzonen-Integration $\sum_{\mathbf{q}}$ erweist sich als aufwendig, denn für $\mathbf{q}\to\mathbf{0}$ divergiert die darin enthaltene Coulomb-Wechselwirkung. Daher

kommt dem Bereich um den Punkt **q**=**0** besonderes Gewicht zu. Für eine konvergente Bestimmung mit Hilfe eines diskreten Abtastverfahrens sind extrem viele **q**-Punkte erforderlich. Allerdings ist das Verhalten der Beiträge für **q**→**0** analytisch bekannt. Dadurch wird es möglich, die Brillouinzonen-Integration mit relativ wenigen **q**-Punkten auszuführen. Anschließend muß eine Korrektur vorgenommen werden, die das divergente Verhalten am Punkt **q**=**0** beschreibt. Hierzu verweisen wir auf Anhang E.

Die abgeschirmte Wechselwirkung $W = \epsilon^{-1}v$ zwischen zwei Ladungen läßt sich formal in zwei Anteile zerlegen:

$$\begin{aligned} W &= v + (\epsilon^{-1} - 1)v \\ &= v + v_{\mathrm{ind}} \quad . \end{aligned}$$

Der erste Beitrag (v) beschreibt die reine Coulomb-Wechselwirkung zwischen zwei Ladungen, während der zweite Term (v_{ind}) die *Abschirmung* beinhaltet: das Feld, das durch die eine Ladung erzeugt wird, wird durch zusätzlich induzierte Ladungen abgeschwächt, so daß letztlich zwischen zwei Ladungen nur eine reduzierte *abgeschirmte Wechselwirkung* zum Tragen kommt. In Analogie dazu bestehen die Eigenwerte $\lambda^{-1}_{\mathbf{q}l}(\omega)$ des Plasmon-Pol-Modells ebenfalls aus zwei Anteilen (siehe Gl. (4.5)). Der erste Term (=1) führt auf eine inverse dielektrische Matrix $\langle\tilde{\epsilon}^{-1}\rangle = \langle 1\rangle$, d. h. dieser Anteil entspricht einem abschirmungsfreien System. Er führt in der Selbstenergie (4.12) zu den Beiträgen (–1) durch die besetzten Zustände n∈Val. Dies entspricht dem Ergebnis der Hartree-Fock-Theorie, in der jegliche Abschirmung unberücksichtigt bleibt. Wir bezeichnen den entsprechenden Anteil der Selbstenergie als Austausch-Anteil Σ_x. Der restliche, frequenzabhängige Term der Eigenwerte $\lambda_{\mathbf{q}l}(\omega)$, der die Parameter $z_{\mathbf{q}l}$ und $\omega_{\mathbf{q}l}$ beinhaltet, führt entsprechend auf den Korrelations-Anteil Σ_c der Selbstenergie. Wir können ihn also als Konsequenz der elektronischen Abschirmeffekte identifizieren.

In der einfachen Form (4.4, 4.5) läßt das Plasmon-Pol-Modell für die inverse dielektrische Matrix eine anschauliche Interpretation der Korrelations-Effekte in der GW-Näherung erkennen. Der frequenzabhängige Anteil kann als *Green-Funktion von Plasmonen* $|\mathbf{q}l\rangle$ aufgefaßt werden. Die Frequenz eines solchen Plasmons ist durch $\omega_{\mathbf{q}l}$ gegeben, und der Vektor $\xi^l_\beta(\mathbf{q})$ erfaßt seine räumlichen Eigenschaften. Zur Veranschaulichung stellen wir in Abb. 4.2 die Plasmon-Dispersion von Si dar, die sich aus dem Plasmon-Pol-Modell ergibt. Die Daten entlang der Hochsymmetrie-Linie Γ—X

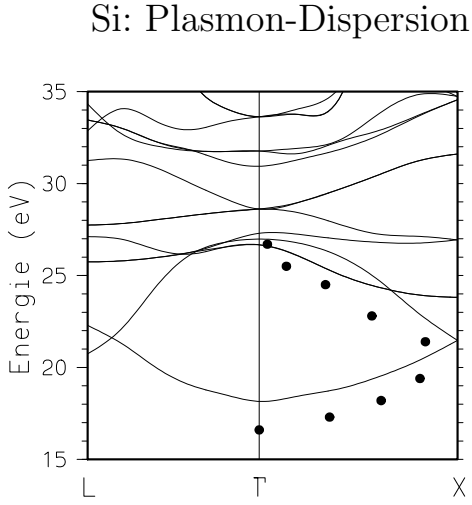


Abbildung 4.2: *Berechnete Plasmon-Dispersion von Si, die sich im Rahmen des Plasmon-Pol-Modells ergibt. Die experimentellen Daten (Punkte) wurden von Chen et al. gemessen [Che80].*

stimmen recht gut mit experimentellen Meßwerten von Chen et al. [Che80] überein. Die Plasmon-Frequenzen stellen nicht nur numerische Parameter eines Modells dar, sondern besitzen auch physikalische Bedeutung. Abb. 4.3 zeigt die räumliche Struktur der zugehörigen Plasmon-Schwingungen mit der niedrigsten Energie am L-Punkt und am X-Punkt. Bei beiden Plasmonen wird der Zusammenhang mit der räumlich inhomogenen Ladungsverteilung deutlich. Die stärkste Dichteänderung findet im Bereich der chemischen Bindungen statt, während der zwischenatomare Bereich wegen der dort geringen Ladungsdichte nur geringe Dichteschwankungen erfährt. Diese Inhomogenität ist bei dem L-Plasmon besonders ausgeprägt, während das X-Plasmon eher einer ebenen Welle gleicht.

Mit dieser Interpretation des Plasmon-Pol-Modells als Green-Funktion von Plasmonen stellt sich der Korrelationsanteil der Selbstenergie im Rahmen der GW-Näherung als *Kopplung zwischen den unkorrelierten Elektronen und den Plasmonen*, ihren eigenen kollektiven Anregungen, dar. Die Kopplung wird durch die Coulomb-Wechselwirkung vermittelt. Es sei an dieser Stelle an die sehr gute Übereinstimmung von GW-Bandstrukturen mit experimentellen Ergebnissen erinnert. Wir stellen daher als ein zentrales Ergebnis fest, *daß diese Kopplung zwischen Elektronen und Plasmonen den wesentlichen Beitrag zur elektronischen Korrelation darstellt.* Weitere Korrelationsmechanismen sind in den von uns betrachteten Systemen von

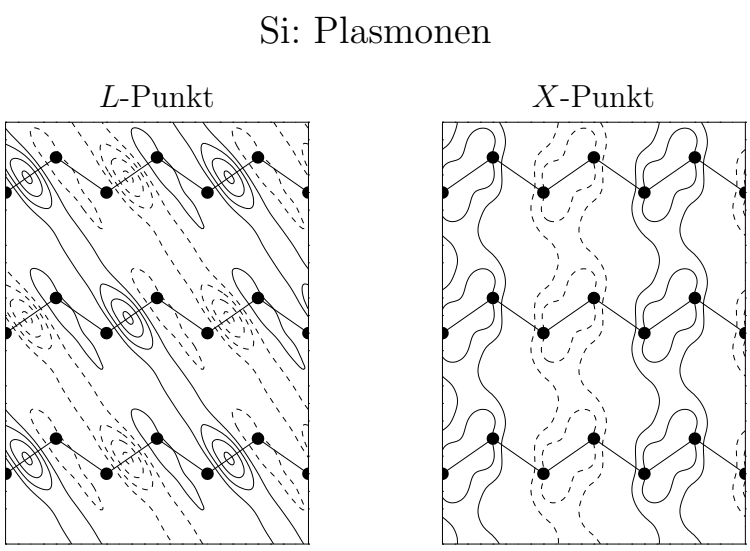


Abbildung 4.3: *Darstellung zweier Plasmonen in Si am L-Punkt (*$\mathbf{q}=(\frac{1}{2},\frac{1}{2},\frac{1}{2})\frac{2\pi}{a}$*; links) und am X-Punkt (*$\mathbf{q}=(1,0,0)\frac{2\pi}{a}$*; rechts). Abgebildet ist der Realteil des Eigenvektors in der* $[\bar{1}10]$*-Ebene. In dieser Momentaufnahme entsprechen gestrichelte Linien einer Abnahme und durchgezogene Linien einer Zunahme an Ladungsdichte.*

untergeordneter Bedeutung.

4.2 Multi-Plasmon-Pol Modelle

Das Plasmon-Pol-Modell bedarf noch einiger Diskussion. Zunächst stellen wir fest, daß die von uns berechneten Bandstrukturen, die aus dem einfachen diagonalen Plasmon-Pol-Modell folgen, mit den Resultaten anderer Autoren sehr gut übereinstimmen, insbesondere mit denen von Godby et al. [God88], die mittels einer direkten Frequenzintegration, also ohne Plasmon-Pol-Modell, erzielt wurden. Dies unterstreicht die Anwendbarkeit einfacher Plasmon-Pol-Modelle zur Berechnung von Quasiteilchen-Energien, zumindest für die typischen IV-, IV-IV- und III-V-Halbleiter in der Nähe des fundamentalen Gaps. Dennoch stellt gerade das sehr einfache Plasmon-Pol-Modell eine massive Näherung dar, die sicher nicht das gesamte Abschirmverhalten erfassen kann. Wir demonstrieren das am Beispiel der makroskopischen dielektrischen Funktion $\epsilon_M(\mathbf{q}=\mathbf{0}, i\omega)$ von Si, die wir in Abb. 4.4 für imaginäre Frequenzen $i\omega$ darstellen. Hier wird deutlich, daß zwischen dem für jede Frequenz einzeln berechneten RPA-Ergebnis und dem Plasmon-Pol-Modell Abweichungen auf-

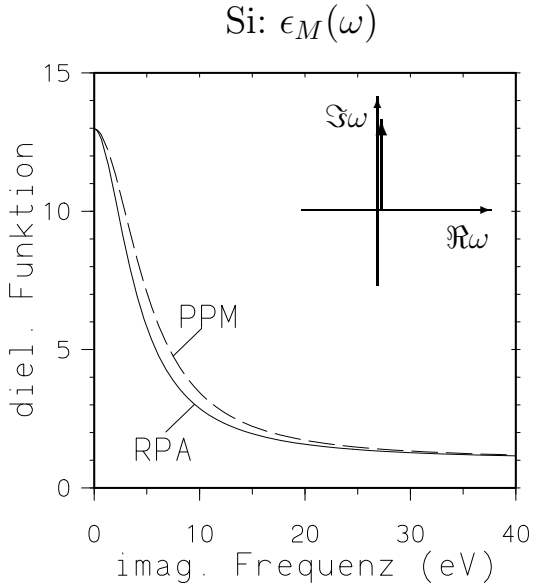


Abbildung 4.4: *Makroskopische dielektrische Funktion $\epsilon_M(\mathbf{q},\omega)$ von Si für $\mathbf{q}$=$\mathbf{0}$, dargestellt für Frequenzen ω auf der imaginären Frequenzachse. Die durchgezogene Kurve zeigt Daten, die im Rahmen der RPA berechnet wurden. Die gestrichelte Kurve zeigt Ergebnisse des einfachen Plasmon-Pol-Modells (4.4).*

treten. [Für $i\omega = 0$ verschwindet die Abweichung natürlich, denn das Plasmon-Pol-Modell wurde bei dieser Frequenz an die RPA-Matrix angepaßt.] Es kann zunächst nicht ausgeschlossen werden, daß auch die Bandstruktur Abweichungen vom exakten Ergebnis zeigt, wenn solch ein einfaches Plasmon-Pol-Modell verwendet wird. Um das zu untersuchen, und um ggf. zu verläßlicheren Quasiteilchen-Bandstrukturen gelangen zu können, haben wir das Plasmon-Pol-Modell erweitert und flexibler gestaltet.

Generell charakterisieren wir ein Plasmon-Pol-Modell (PPM) durch folgende Eigenschaften:

- *global diagonalisierbar/nicht global diagonalisierbar:* unter einem *global diagonalisierbaren* PPM verstehen wir eines, das mittels der Eigenvektoren der statischen dielektrischen Matrix konstruiert wird. Jedem Eigenvektor wird ein frequenzabhängiger Eigenwert zugeordnet, der die Plasmon-Pol-Struktur enthält. In der durch die Eigenvektoren gegebenen Basis ist das Plasmon-Pol-Modell für jede Frequenz, d.h. global, diagonal. Das in (4.4) definierte PPM zählt zu dieser Klasse. Daß die Eigenvektoren der statischen dielektrischen Matrix

(Frequenz ω=0) auch Eigenvektoren der dynamischen dielektrischen Matrix für alle anderen Frequenzen sein sollen, stellt eine prinzipielle Einschränkung der Flexibilität des PPM dar. Diese Einschränkung wird bei der Konstruktion eines *nicht global diagonalisierbaren* PPM nicht verlangt. Bei diesen Modellen wird jedem Matrixelement eine eigene Polstruktur zugeordnet. Hierzu zählt z.B. das von Hybertsen und Louie im Fourierraum definierte Modell [Hyb86]:

$$\tilde{\epsilon}^{-1}_{\mathbf{G},\mathbf{G}'}(\mathbf{q},\omega) = \delta_{\mathbf{G},\mathbf{G}'} + \frac{A_{\mathbf{G},\mathbf{G}'}}{\omega^2 - \omega^2_{\mathbf{G},\mathbf{G}'}(\mathbf{q})} \quad . \tag{4.14}$$

Es soll betont werden, daß die Frage der globalen Diagonalisierbarkeit nichts mit dem Problem zu tun hat, ob eine dielektrische Matrix in der gewählten Funktionenbasis (z.B. in ihrer Fourierdarstellung) diagonal ist oder nicht, ob also lokale Felder vernachlässigt werden oder nicht. Bezüglich dieses Unterscheidungsmerkmals arbeiten wir grundsätzlich mit vollen nichtdiagonalen dielektrischen Matrizen, die die lokalen Felder vollständig berücksichtigen.

- *Anzahl der Pole:* Bei den bisher angeführten PPM wurde jedes relevante Matrixelement mit *einer* Polfrequenz versehen. Es ist ohne weiteres möglich, den einzelnen Pol in (4.5) bzw. in (4.14) durch eine Superposition von n Polen zu ersetzen; in der von uns verwendeten Gauß-Basis läßt sich auf diese Weise z.B. ein (nicht global diagonalisierbares) Modell in der Form

$$\langle\tilde{\epsilon}^{-1}\rangle_{\beta\beta'}(\mathbf{q},\omega) = S_{\beta\beta'}(\mathbf{q}) + \sum_{i=1}^{n} \frac{z^{(i)}_{\beta\beta'}}{\omega^2 - {\omega^{(i)}_{\beta\beta'}}^2} \tag{4.15}$$

konstruieren. Von einem solchen *Multi-Plasmon-Pol-Modell* kann man generell eine höhere Flexibilität zur Darstellung der dynamischen Abschirmeffekte und somit letztlich verläßlichere Bandstruktur-Ergebnisse erwarten.

- *Eingangsdaten:* Um die Parameter eines PPM zu bestimmen, muß Information über das Abschirmverhalten des Systems vorliegen. Man gewinnt sie dadurch, daß die dielektrische Matrix für eine bestimmte Anzahl von (i.a. komplexen) Testfrequenzen im Rahmen der RPA explizit berechnet wird. Es stellt sich heraus, daß für ein n-Pol-Modell die dielektrische Matrix für $2n$ Testfrequenzen bestimmt werden muß, unabhängig davon, ob das PPM global diagonalisierbar sein soll oder nicht. Wenn durch die f-Summenregel (4.10) zusätzliche Information über die Ladungsdichte des Systems einfließt, so reduziert sich die Zahl der Testfrequenzen auf $2n-1$. Für ein einfaches PPM (n=1)

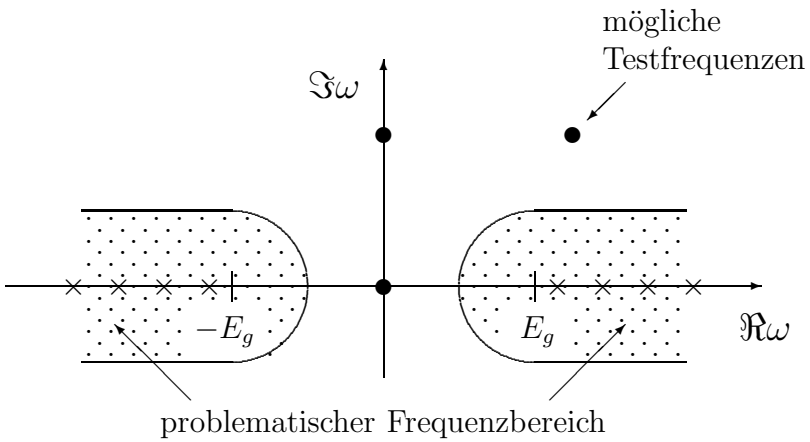


Abbildung 4.5: *Zur konvergenten Berechnung der Polarisierbarkeit (3.12) in RPA für eine Frequenz ω aus dem schattierten Bereich der komplexen Frequenzebene benötigt man ein sehr dichtes* **k***-Punkt-Schema. Deshalb wählen wir die Testfrequenzen zur Festlegung des Plasmon-Pol-Modells außerhalb dieses Bereichs.*

genügt dann bereits die Kenntnis der dielektrischen Matrix für eine Testfrequenz (üblicherweise ω=0) zur Konstruktion des PPM. Allgemein ist bei der Wahl der Testfrequenzen folgendes zu beachten (vgl. Abb. 4.5). Wenn ω in der Nähe einer Übergangsenergie $\Delta E = \pm(E_{\mathrm{m}\mathbf{k}} - E_{\mathrm{n},\mathbf{k}+\mathbf{q}})$ zwischen Valenz- und Leitungszuständen liegt (in der Abbildung durch Kreuze angedeutet), so konvergiert die Polarisierbarkeit (3.12) nur sehr langsam bezüglich der Anzahl der **k**-Punkte, mit denen die erste Brillouin-Zone abgetastet wird. Um also mit wenigen **k**-Punkten die dielektrische Matrix konvergent berechnen zu können, vermeide man Testfrequenzen aus dem in Abb. 4.5 schattiert dargestellten Bereich. Zur Ermittlung der Modellparameter aus den explizit berechneten Matrizen verweisen wir auf Anhang F.

In Tabelle 4.1 stellen wir diejenigen PPM zusammen, die in dieser Arbeit Verwendung finden. Am meisten benutzen wir das PPM I, da es den geringsten numerischen Aufwand erfordert (global diagonalisierbare Form; nur eine Testfrequenz) und gleichzeitig sehr verläßliche Bandstrukturergebnisse im Gap-Bereich erzielt, wie wir unten zeigen werden.

Generell besitzen die Multi-PPM mit mehreren Polen eine höhere Flexibilität als die einfachen PPM und ermöglichen somit eine verläßlichere Reproduktion der tatsächlichen dielektrischen Funktion. Wir demonstrieren das in Abb. 4.6, in der wir die Ergebnisse zweier Modelle mit der in RPA berechneten makroskopischen

	global diagonalisierbar	Zahl n der Pole	f-Summen-Regel	Zahl der Testfrequenzen	Bemerkungen
I	×	1	×	1 ($\omega = 0$)	vgl. [Lin88]
II		1	×	1 ($\omega = 0$)	vgl. [Hyb86]
III	×	1		2 ($\omega = 0, i\omega_0$)	
IV		$n = 2, 3$	×	$2n - 1$	Multi-PPM

Tabelle 4.1: *Zusammenstellung der in dieser Arbeit untersuchten und verwendeten Plasmon-Pol-Modelle.*

dielektrischen Funktion für **q=0** vergleichen. Die Darstellung erfolgt entlang zweier Wege in der komplexen Frequenzebene, die die schattierten Bereiche aus Abb. 4.5 vermeiden. Aus der Abbildung wird deutlich, daß die Resultate des einfachen PPM I erheblich von den exakten RPA-Ergebnissen abweichen, während das Multi-PPM IV mit n=2 bzw. 3 Polen die Frequenzabhängigkeit der dielektrischen Funktion sehr zuverlässig wiedergibt.

Die sehr gute Übereinstimmung des Multi-PPM IV mit der RPA-Funktion läßt erwarten, daß mit diesem Modell sehr zuverlässige GW Quasiteilchen-Bandstrukturen berechnet werden können. Wir vergleichen daher in Abb. 4.7 die Bandstruktur-Resultate der verschiedenen PPM für Si miteinander, wobei letztlich der Vergleich mit dem Multi-PPM IV ausschlaggebend für die Beurteilung der Zuverlässigkeit und Leistungsfähigkeit eines jeden einzelnen PPM ist. Als Testfrequenzen für das PPM III wählen wir die Werte $\omega_1 = 0$ (statische dielektrische Matrix) sowie $\omega_2 = i \cdot 1$ Ryd (auf der imaginären Frequenzachse). Das Multi-PPM IV untersuchen wir sowohl mit $n = 2$ Polen (IVa) als auch mit $n = 3$ Polen (IVb). Als Testfrequenzen wählen wir die Werte 0, $i \cdot 1$ Ryd und $(1 + i)$ Ryd für das Modell IVa. Für IVb nehmen wir noch die Frequenzen $i \cdot 2$ Ryd und $(2 + i)$ Ryd hinzu. Die Modelle erweisen sich allgemein als sehr stabil gegenüber der Wahl der Testfrequenzen, solange diese nicht aus dem verbotenen Bereich stammen (siehe oben) und sich deutlich genug unterscheiden (sonst treten Instabilitäten bei der Ermittlung der Plasmon-Frequenzen und -gewichte auf). Die in Abb. 4.7 dargestellten Bandstruktur-Ergebnisse erlauben folgendes Fazit:

- Die Ergebnisse, die mit den PPM I und II ermittelt wurden, stimmen her-

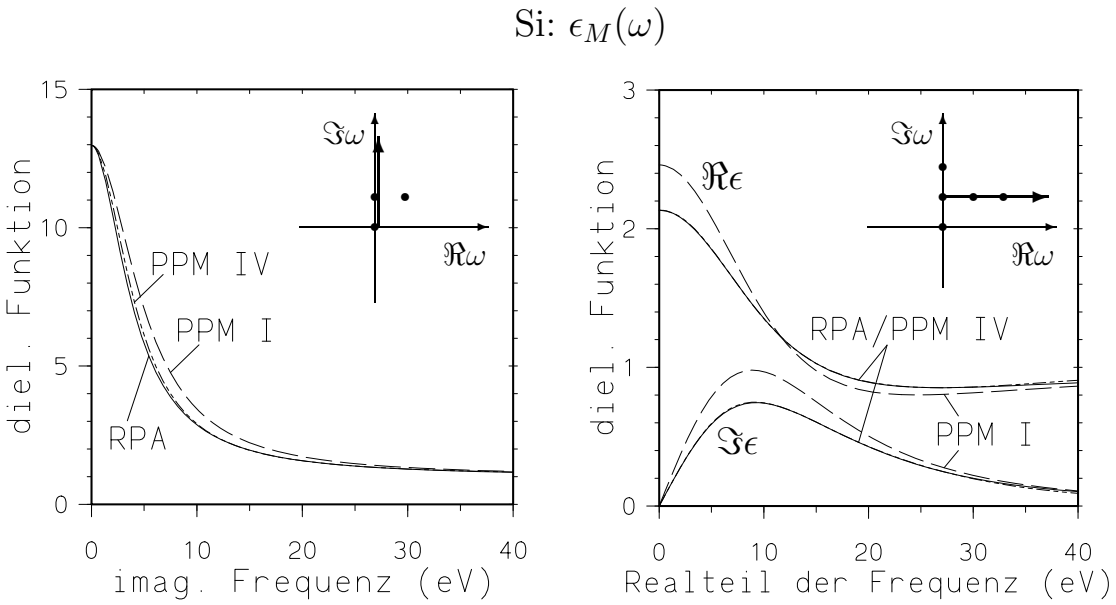


Abbildung 4.6: *Makroskopische dielektrische Funktion $\epsilon_M(\mathbf{q},\omega)$ von Si für $\mathbf{q}$=**0**, dargestellt für Frequenzen ω auf der imaginären Frequenzachse (links) und auf einem Weg parallel zur reellen Achse (rechts; der Imaginärteil der Frequenzen beträgt 1 Ryd = 13.6 eV). Die durchgezogene Kurve zeigt Daten, die im Rahmen der RPA berechnet wurden. Die gestrichelte Kurve zeigt Ergebnisse des PPM I, die lang-kurz gestrichelte Kurve Ergebnisse des PPM IV. Für das PPM IV wurden in der linken Abbildung die Testfrquenzen 0, i·1 Ryd und (1+i) Ryd verwendet (im Koordinatensystem durch Punte angedeutet). Für die rechte Abbildung wurden zusätzlich die Testfrequenzen i·2 Ryd und (2+i) Ryd verwendet. In der rechten Figur stimmen die Ergebnisse des PPM IV fast exakt mit der RPA überein.*

vorragend miteinander überein. Ob ein PPM global diagonalisierbar ist oder nicht, spielt für die Berechnung von GW Quasiteilchen-Bandstrukturen keine Rolle.

- Im unteren Valenzbandbereich und im oberen Leitungsbandbereich ergeben sich leichte Differenzen von bis zu 0.5 eV zwischen den PPM I, II einerseits und den PPM III, IV andererseits. Wir schließen daraus, daß es zu Ungenauigkeiten in der Bandstruktur führen kann, wenn nur die statische dielektrische Matrix berücksichtigt wird (PPM I und II) und keine Information über weitere Testfrequenzen in das Modell eingeht. Die Berücksichtigung der f-Summenregel in den PPM I und II vermag dieses Defizit nicht vollständig zu kompensieren.

- Im Bereich des fundamentalen Gaps, also im Energieintervall von -5 eV bis +5

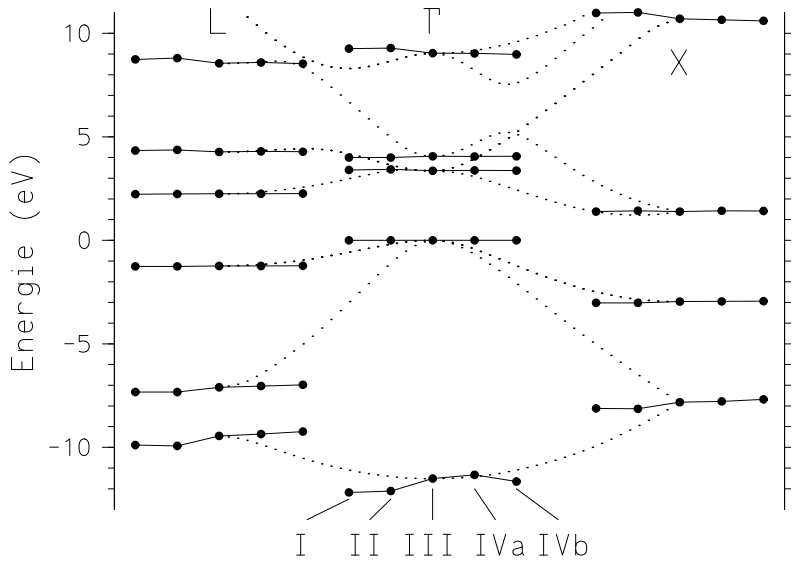


Abbildung 4.7: *GW Quasiteilchen Bandstruktur-Energien für Si, mit den verschiedenen Plasmon-Pol-Modellen I bis IV berechnet (siehe Text).*

eV, führen alle PPM auf die gleiche Bandstruktur. Wir stellen als zentrales Ergebnis der Untersuchungen der PPM fest, *daß man schon mit einem sehr einfach strukturierten PPM sehr genaue Bandstrukturergebnisse erzielen kann, wenn man primär an Zuständen in diesem Bereich in der Nähe des Gaps interessiert ist.* Die Untersuchungen an Halbleiteroberflächen, die wir in Teil II der Arbeit vorstellen, führen wir grundsätzlich mit dem PPM I durch.

Kapitel 5

Eigenschaften der Selbstenergie

In diesem Abschnitt sollen die verschiedenen Beiträge zur Selbstenergie näher betrachtet werden. Wir hatten im letzten Abschnitt die Selbstenergie in zwei Anteile zerlegt:

$$\Sigma = \Sigma_x + \Sigma_c \quad . \tag{5.1}$$

Der Anteil Σ_x beschreibt die Austauscheffekte eines unkorrelierten Elektronensystems. Er stellt sich letztlich als Konsequenz des Pauli-Prinzips dar, das sich aus der Fermi-Statistik für die Elektronen ergibt. Der Korrelations-Anteil Σ_c hingegen vereinigt alle Folgen der Abschirmung der Coulomb-Wechselwirkung in sich, inklusive dynamischer Effekte und lokaler Felder. Die Aufteilung (5.1) ist unter anderem von Interesse, weil sie eine Zerlegung in einen relativ einfach zu behandelnden Term (Σ_x) und einen aufwendigen Anteil bedeutet. Bei einer konkreten Rechnung erlaubt der Austausch-Anteil eine zuverlässige Kontrolle über das System; unter anderem ermöglicht er eine Beurteilung der Konvergenzeigenschaften der Basissätze. Der Korrelationsanteil hingegen macht den eigentlichen Reiz des Formalismus aus. Im Rahmen der GW-Näherung beinhaltet der Korrelationsanteil seinerseits weitere Approximationen, wie etwa das oben erwähnte Plasmon-Pol-Modell. Er stellt einen Ansatzpunkt für Erweiterungen der reinen GW-Näherung dar, falls sich diese bei stärker korrelierten Systemen als unzureichend erweisen sollte.

In Abb. 5.1 stellen wir am Beispiel von Si die Erwartungswerte des Austauschenergie-Operators Σ_x, des Korrelationsenergie-Operators Σ_c und des gesamten Selbstenergie-Operators Σ dar, und zwar in Abhängigkeit von der jeweiligen LDA-Energie der betrachteten Zustände. Wir vergleichen die Daten mit entsprechenden Werten für das homogene Elektronengas gleicher mittlerer Dichte r_s=2 (durchgezogene Lini-

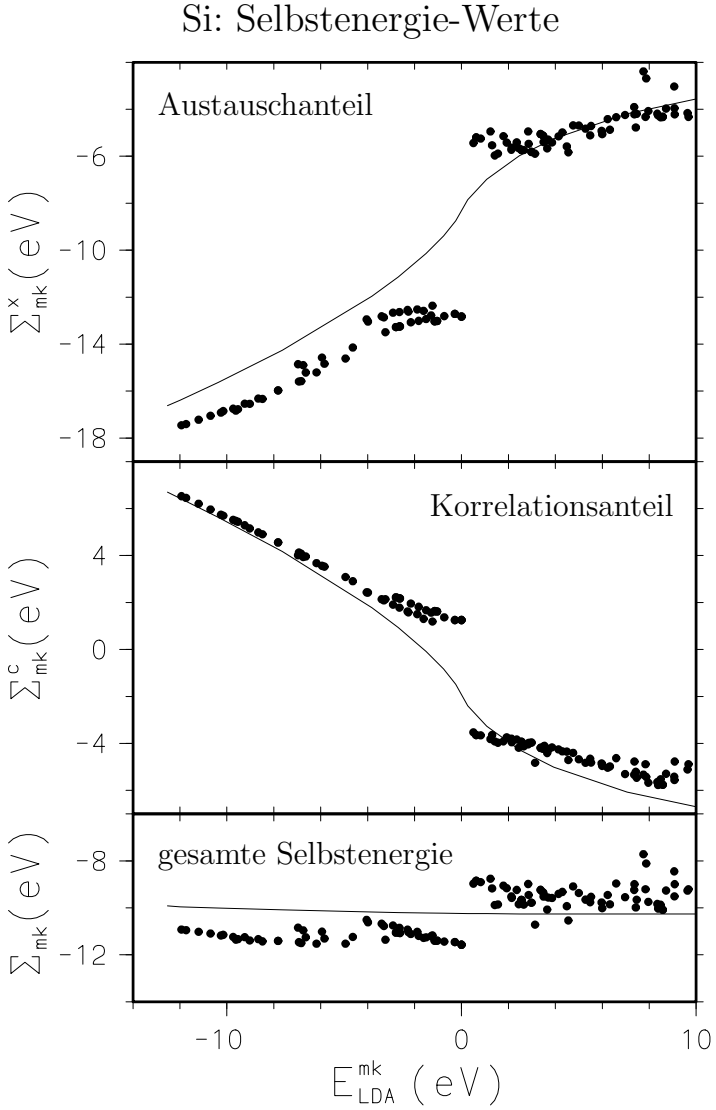


Abbildung 5.1: *Selbstenergiewerte elektronischer Zustände. Die Datenpunkte beziehen sich auf Si. Die durchgezogenen Kurven beziehen sich auf ein homogenes Elektronengas mit der gleichen mittleren Dichte wie Si (r_s=2) [Hed69]. Der Austauschanteil (oben), der Korrelationsanteil (Mitte) und die gesamte Selbstenergie (unten) werden mit der LDA-Energie des jeweiligen Zustandes verglichen. Das Valenzband-Maximum bzw. die Fermienergie wurden auf 0 eV festgelegt.*

en) [Hed69]. [Um für Si diese mittlere Dichte zu realisieren, wurden die Rechnungen bei der Gitterkonstante a_0=5.416 Å durchgeführt.] Für die grundsätzlich negative Austausch-Energie stellen wir eine starke Abhängigkeit von der Energie des jeweiligen Zustandes fest. Je höher $E^{\mathrm{LDA}}_{\mathbf{mk}}$ ist, desto höher ist auch die Austausch-Energie; Σ_x wächst monoton. Das gilt sowohl für das homogene Elektronengas als auch für

den inhomogenen Halbleiter. Die Daten der beiden Systeme stimmen im großen und ganzen miteinander überein, unterscheiden sich aber in zwei wesentlichen Punkten:

- Die Austausch-Energien der Valenzband-Zustände ($E_{\mathrm{m}\mathbf{k}} < 0$) des Halbleiters liegen um ca. 2 eV unter denen des homogenen Elektronengases. Das könnte darauf zurückzuführen sein, daß diese Zustände, die ja besetzt werden und somit für die elektronische Ladung verantwortlich sind, eine Ladungsdichte verspüren, die die mittlere Dichte des Halbleiters deutlich übersteigt. Eine größere Dichte (kleineres r_s) führt generell zu tieferen Austauschenergien.

- In der Nähe des fundamentalen Gaps ($E_{\mathrm{m}\mathbf{k}} = 0$) erfahren die besetzten Valenzband-Zustände des Halbleiters eine zusätzliche Verschiebung ihrer Austauschenergie zu tieferen Energien, die unbesetzten Leitungsband-Zustände hingegen zu höheren Energien. Dadurch ergibt sich eine übermäßig große „Austausch-Energielücke“ von 7 eV zwischen Valenz- und Leitungsband-Zuständen. Dies führt letztlich im Falle von Halbleitern zu schlechter Übereinstimmung zwischen korrelationsfreien Hartree-Fock-Bandstrukturen und experimentellen Daten, insbesondere zu einer deutlichen Überschätzung der fundamentalen Energielücke um 5 eV im konkreten Beispiel von Si.

Der Korrelationsanteil der Selbstenergie zeigt eine entgegengesetzte Energieabhängigkeit. Während die Austauschenergie mit steigender Energie $E_{\mathrm{m}\mathbf{k}}$ ansteigt, fällt der Korrelationsanteil mit wachsender Energie $E_{\mathrm{m}\mathbf{k}}$. Er ist für besetzte Zustände eines Halbleiters in der Regel positiv und für unbesetzte Zustände meistens negativ. Im großen und ganzen ergeben sich für die Zustände des Halbleiters Korrelationsenergien, die mit denen des homogenen Elektronengases übereinstimmen. Ein wesentlicher Unterschied besteht auch hier in der Nähe des fundamentalen Gaps. Die besetzten Valenz-Zustände des Halbleiters erfahren eine größere Korrelations-Energie als im homogenen Elektronengas, während die Korrelations-Energie der unbesetzten Leitungs-Zustände erniedrigt wird. Das bewirkt eine „Korrelations-Energielücke“ von ca. –5 eV zwischen besetzten und unbesetzten Zuständen. Hierdurch wird das zu große Hartree-Fock-Ergebnis für die fundamentale Energielücke um 5 eV reduziert.

Für die gesamte Selbstenergie ergibt sich folgendes Verhalten. Die Selbstenergie des homogenen Elektronengases ist über einen weiten Energiebereich nahezu konstant; zwischen –12 eV und +10 eV sinkt sie nur um 0.3 eV. Die logarithmischen

Energie-Abhängigkeit des Selbstenergie-Operators:

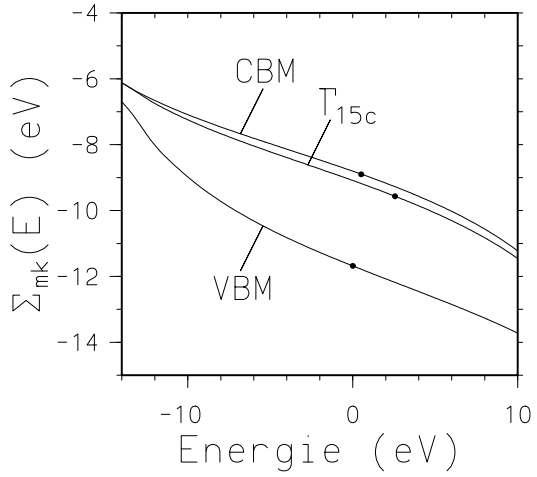


Abbildung 5.2: *Explizite Energieabhängigkeit der Selbstenergie, d.h. des Erwartungswerts des Selbstenergieoperators $\Sigma(E)$. Die Kurven beziehen sich auf die elektronischen Zustände in Si am Valenzband-Maximum (VBM) und am Leitungsband-Minimum (CBM) sowie auf den tiefsten Leitungsband-Zustand am Γ-Punkt. Die Punkte deuten die jeweilige LDA-Energie $E^{\mathrm{LDA}}_{\mathbf{mk}}$ an. Das VBM hat die Energie 0 eV.*

Divergenzen, die sich sowohl in der Energieableitung der Austauschenergie als auch in der des Korrelationsanteils jeweils an der Fermikante finden, heben sich gegenseitig auf. Daß die Selbstenergie nahezu konstant ist, führt dazu, daß homogene Elektronensysteme und einfache Metalle in guter Näherung als effektiv wechselwirkungsfreie Elektronensysteme behandelt werden können. Die Selbstenergie des Halbleiters Si weicht von diesem Verhalten ab. Die besetzten Zustände erfahren generell eine abgesenkte und die unbesetzten Zustände eine angehobene Selbstenergie. Das Selbstenergie-Gap am Valenzband-Maximum beträgt 2.5 eV. Daraus resultiert letztlich die Existenz einer fundamentalen Energielücke in der Bandstruktur des Halbleiters. Wir merken an, daß die Erwartungswerte des LDA-Austausch-Korrelations-Potentials für den Halbleiter ein ähnliches Verhalten wie die Selbstenergien aufweisen. Daraus ergibt sich die oft recht gute Übereinstimmung zwischen den Banddispersionen in GW Quasiteilchen-Bandstrukturen und LDA-Bandstrukturen. Lediglich die Größe des Selbstenergie-Gaps fällt geringer aus. Daher wird die fundamentale Energielücke zwischen den Valenz- und Leitungsbändern von der LDA unterschätzt.

In diesem Zusammenhang sei daran erinnert, daß die Selbstenergie-Werte in zwei-

erlei Weise von der Energie abhängen. Zum einen erfahren verschiedene Zustände (mit verschiedenen Energien $E_{\mathrm{m}\mathbf{k}}$) natürlich auch unterschiedliche Selbstenergien $\Sigma_{\mathrm{m}\mathbf{k}}$. Wenn die entsprechenden Wertepaare (Energie und Selbstenergie eines jeden Zustandes) zusammengestellt werden, so ergibt sich eine implizite Energieabhängigkeit der Selbstenergie, wie wir sie in Abb. 5.1 diskutiert haben. Zum anderen hängt der Selbstenergie-Operator $\Sigma(E) = \Sigma_x + \Sigma_c(E)$ in seinem Korrelationsanteil explizit von der Energievariable E ab. Das führt dazu, daß die Selbstenergie $\langle \mathrm{m}\mathbf{k}|\Sigma(E)|\mathrm{m}\mathbf{k}\rangle$ eines bestimmten Zustandes m$\mathbf{k}$ über einen weiten Energiebereich (zumindest in der Nähe der zugehörigen Quasiteilchen-Energie) mit wachsender Energie E monoton fällt (siehe Abb. 5.2). Allerdings ist die Energie E letztlich keine freie Variable, sondern wird über die Dyson-Gleichung $E_{\mathrm{m}\mathbf{k}} = E^0_{\mathrm{m}\mathbf{k}} + \Sigma_{\mathrm{m}\mathbf{k}}(E_{\mathrm{m}\mathbf{k}})$ festgelegt. Auf diese explizite Energieabhängigkeit des Selbstenergie-Operators kommen wir in Abschnitt 14 zurück.

Teil II

Bandstrukturen von Halbleiteroberflächen

Oberflächen von Festkörpern spielen aus zwei Gründen eine wichtige Rolle in der Festkörperphysik und bei technologischen Fragestellungen. Einerseits beeinflussen die Oberflächeneffekte jeden Einblick in den Volumenfestkörper, denn eine experimentelle Ankopplung an das Volumensystem kann nur durch die Oberfläche hindurch erfolgen. Wenn man primär an den reinen Volumeneigenschaften interessiert ist, so ist es unabdingbar, die Eigenschaften und Einflüsse der Oberfläche zu erkennen, um sie aus der Analyse ausschließen zu können. Andererseits stellt eine Oberfläche selber ein hochinteressantes physikalisches System dar. Die um eine Dimension reduzierte Translationssymmetrie und die gegenüber dem Volumensystem geänderte chemische Umgebung der Oberflächenatome führen zum Auftreten völlig neuer Eigenschaften und Effekte. Die Oberfläche eines Kristalls hat zwar gegenüber dem wesentlich größeren, volumenartigen Gesamtsystem nur einen geringen Raumanteil. Wenn aber der Kristall von außen betrachtet wird, so ist es zuallererst die Oberfläche, die wahrgenommen wird. Chemische Prozesse wie Adsorption und Desorption von Adatomen und Molekülen spielen sich unmittelbar auf atomarer Längenskala an der Oberfläche ab. Auch die Wechselwirkung anderer, insbesondere geladener Teilchen mit dem Kristall finden oftmals hauptsächlich an der Oberfläche statt. Bei denjenigen Experimenten, die mit unseren theoretischen Untersuchungen in Zusammenhang stehen, spielen Elektronen eine wichtige Rolle. Sie besitzen in einem Energiebereich von 40-100 eV eine freie Weglänge von nur wenigen Å und können dann nicht tiefer als einige Atomschichten in das Material eindringen [Hen94]. Elektronen des gleichen Energiebereichs, die aus dem System austreten, stammen vorwiegend aus dem oberflächennahen Bereich, denn Elektronen aus tieferen Schichten besitzen wegen der geringen freien Weglänge eine große Streuwahrscheinlichkeit und können kaum ohne inelastische Stoßprozesse das System verlassen. Hieraus leitet sich eine große Oberflächensensitivität aller Experimente ab, an denen Elektronen dieses Energiebereichs beteiligt sind.

Die Besonderheiten der Oberfläche gegenüber dem Volumenfestkörper lassen sich in zwei Klassen einteilen.

(i) Die *geometrische Anordnung* der Atome unterscheidet sich von der idealen Struktur, die man aus dem Volumenkristall erhält, wenn auf der einen Seite einer Trennebene der halbe Kristall entfernt wird. An der so definierten Oberfläche kommt es infolge des Fehlens chemischer Bindungspartner zu einer Umverteilung elektronischer Ladung und daraufhin zu einer Verschiebung der Atome in neue Gleichgewichtspositionen. Die treibende Kraft einer solchen *Relaxation* sowie von Rekonstruktionen

ist der damit erzielbare Gewinn an Gesamtenergie. Eine merkliche Relaxation findet meist nur in den oberflächennahen Schichten bis zu einer Tiefe von ca. 5 Å statt. Die Verschiebungen der Atome relativ zu ihren idealen Positionen betragen in der Regel wesentlich weniger als 1 Å. Größere Verschiebungen treten auf, wenn durch das Ausbilden neuer chemischer Bindungen die Gesamtenergie erheblich verringert werden kann. Bei der Dimerbildung an den Oberflächen C(001), Si(001) und Ge(001) kommt es zu Relaxationen von mehr als 1 Å.
In vielen Fällen weist das Verschiebungsmuster eine geringere zweidimensionale Translationssymmetrie auf, als durch das zugrunde liegende Volumensystem vorgegeben wird. Dieses Ausbilden einer Überstruktur mit einer vergrößerten Oberflächen-Einheitszelle wird als *Rekonstruktion* bezeichnet. Wenn auch die Untersuchung der geometrischen Atomstruktur nicht selbst Gegenstand dieser Arbeit ist, so spielt doch die Geometrie eine entscheidende Rolle, denn sie steht in direktem Zusammenhang mit den elektronischen Eigenschaften. Die elektronischen Fragestellungen können nur auf der Basis einer genauen Kenntnis der Atomstruktur zuverlässig behandelt werden.
(ii) Zusätzlich zu den Volumenzuständen bilden sich an der Oberfläche des Kristalls weitere elektronische Zustände aus. Diese *Oberflächenzustände* treten besonders markant in Erscheinung, wenn sie sich in ihrer Energie $E_{\mathrm{m}\mathbf{k}_\parallel}$ von den Volumenzuständen unterscheiden, d.h. wenn sie im Gap oder in einer Tasche der projizierten Volumenbandstruktur liegen. Sie können dann nicht resonant mit den Volumenzuständen koppeln und sind auf den Oberflächenbereich beschränkt. Ihre räumliche Aufenthaltswahrscheinlichkeit nimmt in den Festkörper hinein exponentiell ab. In oberflächensensitiven experimentellen Untersuchungen treten sie besonders deutlich hervor. Solche Zustände werden wir in den Mittelpunkt dieses Abschnitts stellen. Ein weiterer elektronischer Effekt an der Oberfläche betrifft die Core-Elektronen des Substratmaterials. Infolge der Umverteilungen der elektronischen Valenzladungsdichte weicht das Kristallpotential an den Oberflächen-Atomkernen vom Wert im Volumenkristall ab. Außerdem liegen an der Oberfläche andere Abschirmeigenschaften und somit andere Korrelationsenergien als im Volumen vor. Beides führt dazu, daß sich die Bindungsenergie der Core-Elektronen an diesen Atomen vom entsprechenden Volumenwert unterscheidet. Diese *chemische Verschiebung* der Core-Niveaus kann bis ± 1 eV betragen. Sie findet nur in wenigen Atomschichten an der Oberfläche in meßbarer Stärke statt. Wir werden solche Verschiebungen in Kap. 15 im Zusammenhang mit der Untersuchung der Coreniveaus betrachten.

Kapitel 6

Das Slabkonzept

In einem kristallinen Volumenfestkörper liegt in jeder Raumrichtung diskrete Translationssymmetrie vor. Das System besitzt eine dementsprechend kleine dreidimensionale Einheitszelle. Bei einem Oberflächensystem ist die Translationssymmetrie senkrecht zur Oberfläche, also in der z-Richtung, nicht mehr vorhanden. Bezüglich dieser Richtung lassen sich die Zustände nicht mehr nach dem Bloch-Theorem klassifizieren; die minimale Einheitszelle ist unendlich lang. Die theoretische Beschreibung des Systems wird dadurch wesentlich erschwert. Die Standardmethoden der Volumenkristallphysik sind zunächst nicht anwendbar. Es bestehen mehrere Möglichkeiten, zu einer Beschreibung von Oberflächen zu kommen [Pol80]. In dieser Arbeit beschreiben wir die Oberfläche mit Hilfe einer Kombination aus *Slab- und Superzellenmethode* [Sch75]. Dabei wird ein Stapel von mehreren Atomschichten (ein *Slab*) betrachtet, der beidseitig von Vakuum begrenzt wird. Er weist konstruktionsbedingt *zwei* Oberflächen auf, die wir der Anschauung halber als obere und untere Oberfläche (bzw. als Ober- und Unterseite) bezeichnen werden. Dieses Schichtsystem aus Slab und Vakuum wird in z-Richtung periodisch wiederholt. Dadurch wird künstlich eine diskrete Translationsinvarianz auch in dieser Richtung erzeugt, und die Standardverfahren der Volumenfestkörperphysik können zur Anwendung gelangen. Es muß eine entsprechend große Einheitszelle (eine sogenannte Superzelle) eingeführt werden, die in z-Richtung um ein Vielfaches länger ist als die Volumen-Einheitszelle des Substrats. Der Aufbau eines Slabsystems, insbesondere die Dicke des Slabs und des Vakuums, orientiert sich an folgenden Anforderungen (vgl. Abb. 6.1).
(i) Wenn der Verlauf einer physikalischen Größe (etwa der Ladungsdichteverteilung oder des Potentials), von der Oberfläche beginnend, in den Festkörper hinein verfolgt wird, so stellt man fest, daß ihr Verhalten immer mehr dem des Volumenfestkörpers

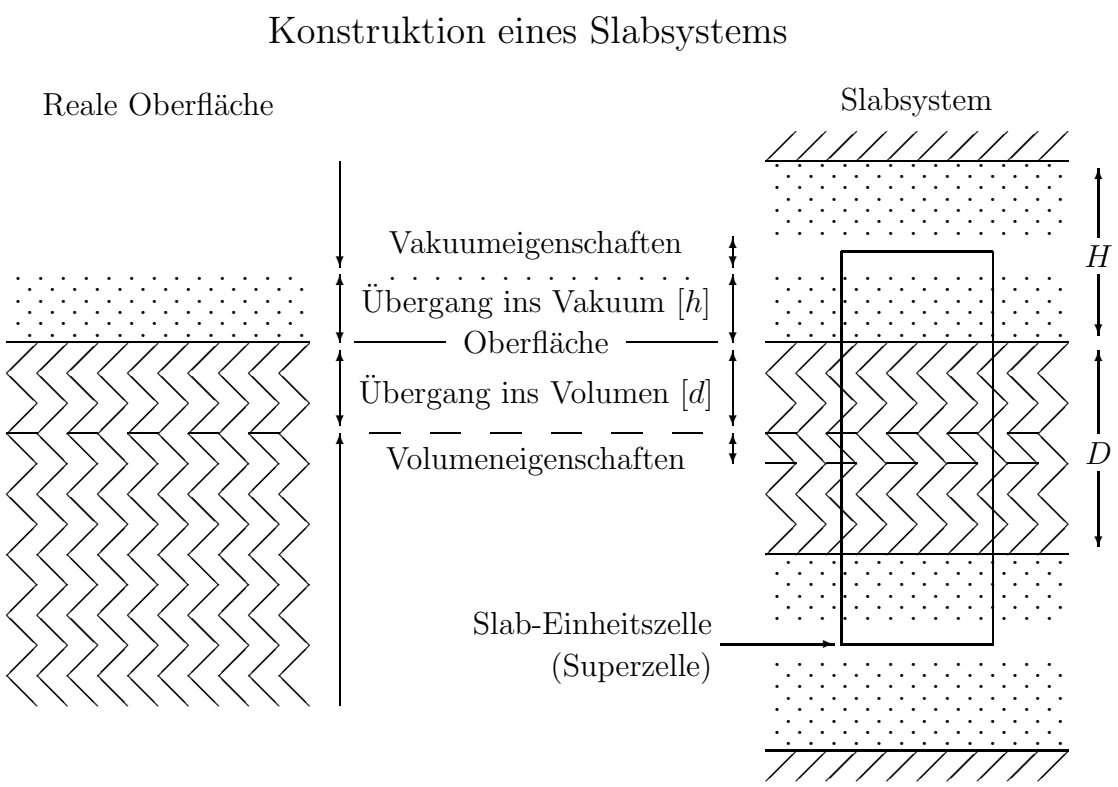


Abbildung 6.1: *Konstruktion eines Slabsystems (rechts), das zur Simulation eines realen Oberflächensystems (links) dient. Der schraffierte Bereich stellt die Atomschichten dar. Weitere Erläuterungen siehe Text.*

sebst entspricht. In einer (materialabhängigen) Tiefe d ist das System volumenartig, und die Oberfläche ist nicht mehr verspürbar. Ein Slab, der die Oberfläche simulieren soll, muß so dick sein, daß er beidseitig den Übergang zum Volumencharakter ermöglicht. Dazu muß seine Dicke D größer als $2d$ sein.
(ii) Von der Oberfläche in das Vakuum hinein nimmt die elektronische Ladungsdichte exponentiell ab, und das Potential nähert sich asymptotisch seinem Vakuumwert V_{Vak}=0. Ab einer Höhe h liegen reine Vakuumeigenschaften vor. Für das Slabsystem bedeutet das, daß beiderseits des Slabs mindestens ein Vakuum der Dicke h vorliegen muß. Die periodisch wiederholten Slabs müssen mit einer Vakuumschicht voneinander getrennt werden, deren Dicke mindestens $H = 2h$ beträgt.

Bevor wir diese Kriterien am Beispiel der Si(001)-Oberfläche genauer untersuchen, soll noch zwischen den Begriffen Slabmethode und Superzellenmethode differenziert werden. Der Begriff „Superzellenmethode“ orientiert sich an der Beschreibung des Systems mittels sich periodisch wiederholender Slabs. Die Wechselwir-

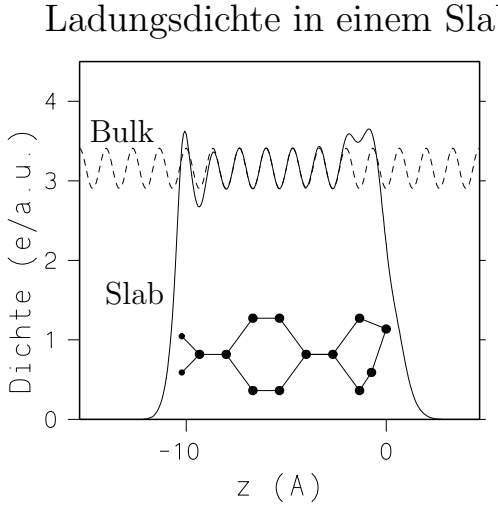


Abbildung 6.2: *Ladungsdichteverlauf des Si(001)-(2×1) Slabsystems entlang der z-Achse im Vergleich zum entsprechenden Verlauf im Volumenkristall. Die Dichte wurde bezüglich der x- und y-Richtung integriert. Das Slabsystem besteht aus acht Schichten Si. Auf der Oberseite (rechts im Bild; $z = 0\mathring{A}$) befinden sich asymmetrische Dimere. Auf der Unterseite (links) sind die Dangling Bonds mit H-Atomen abgesättigt. Die Darstellung erfolgt über genau eine Superzelle.*

kung über das Vakuum hinweg wird dabei zunächst nicht näher spezifiziert. In einer konkreten Rechnung muß das Vakuum so groß gewählt werden, daß die prinzipiell mögliche Wechselwirkung real Null wird. Hingegen besteht das System im Rahmen einer „Slabmethode“ nur aus einem einzigen Slab mit beidseitig halbunendlichem Vakuum. Eine Wechselwirkung mit anderen Slabs ist gar nicht möglich. In unserem Zugang finden sich Elemente beider Vorgehensweisen. Der Hamilton-Operator wird mit lokalisierten Gauß-Orbitalen dargestellt, und zwar sowohl in der LDA als auch – unter Hinzunahme des Selbstenergieoperators – in der GWA. In dieser Basis lassen sich Wechselwirkungen zwischen den Slabs über das Vakuum hinweg exakt ausschließen; der Hamiltonian beschränkt sich auf genau einen Slab. Die Darstellung der Potentiale erfolgt hingegen als Fourierreihe. Durch die – auch in z-Richtung – diskrete Rasterung im reziproken Raum wird im reellen Raum automatisch eine periodische Wiederholung des Slab-Vakuum-Systems erzeugt. Das ermöglicht grundsätzlich eine Wechselwirkung des Slabs mit seinem „Spiegelbild“ über das Vakuum hinweg. Unser Zugang stellt also eine Kombination aus Slab- und Superzellenmethode dar.

Wir weisen nun am Beispiel der Si(001)-(2×1) Oberfläche nach, daß die oben gestellten Anforderungen erfüllt sind. Der von uns verwendete Slab umfaßt acht

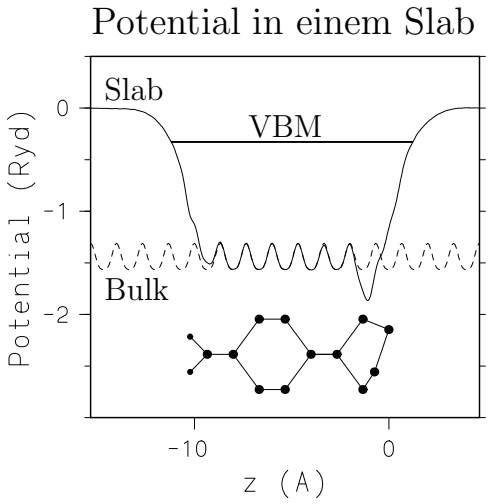


Abbildung 6.3: *Potentialverlauf (LDA; ohne nichtlokalen Pseudopotentialanteil) des Si(001)-(2×1) Slabsystems entlang der z-Achse im Vergleich zum entsprechenden Verlauf im Volumenkristall. Das Potential wurde bezüglich der x- und y-Richtung gemittelt. Die waagerechte Linie gibt die Energie des Volumen-Valenzbandmaximums (VBM) an, auf das wir im Folgenden alle Bandstrukturenergien beziehen werden.*

Schichten Si und eine Schicht H auf der Unterseite [den Zweck einer chemisch unterschiedlichen Behandlung der beiden Slaboberflächen diskutieren wir weiter unten]. Die Dicke des Slabs beträgt 10.3 Å. Der Vakuumabstand zum nächsten Slab beträgt 9.8 Å. In Abb. 6.2 stellen wir die in x- und y-Richtung integrierte Ladungsdichte dar (durchgezogene Kurve). Sie stimmt in der Slabmitte sehr gut mit der entsprechenden integrierten Ladungsdichte des Si Volumenkristalls (gestrichelt) überein. An den Oberflächen ergeben sich deutliche Unterschiede zum Volumenkristall. Im Vakuum klingt die Ladungsdichte sehr schnell auf Null ab. Abb 6.3 zeigt in entsprechender Weise das gemittelte Gesamtpotential (LDA) entlang der z-Achse. Die Energieskala wurde in dieser Darstellung so gewählt, daß das Potential im Vakuum den Wert Null annimmt. Hinsichtlich des räumlichen Verlaufs des Potentials stellt man ebenfalls in der Slabmitte eine sehr gute Übereinstimmung mit dem Volumensystem fest. In das Vakuum hinein konvergiert das Slabpotential schnell gegen seinen asymptotischen Grenzwert. Bezüglich der Ladungsdichte und des Potentials sind die Anforderungen an die Konstruktion des Slabs erfüllt; der Slab und das umgebende Vakuum sind dick genug gewählt, und die Eigenschaften der Slaboberfläche werden denen der entsprechenden Oberfläche eines halbunendlichen Kristalls entsprechen.

Lokale Zustandsdichten eines Slabsystems

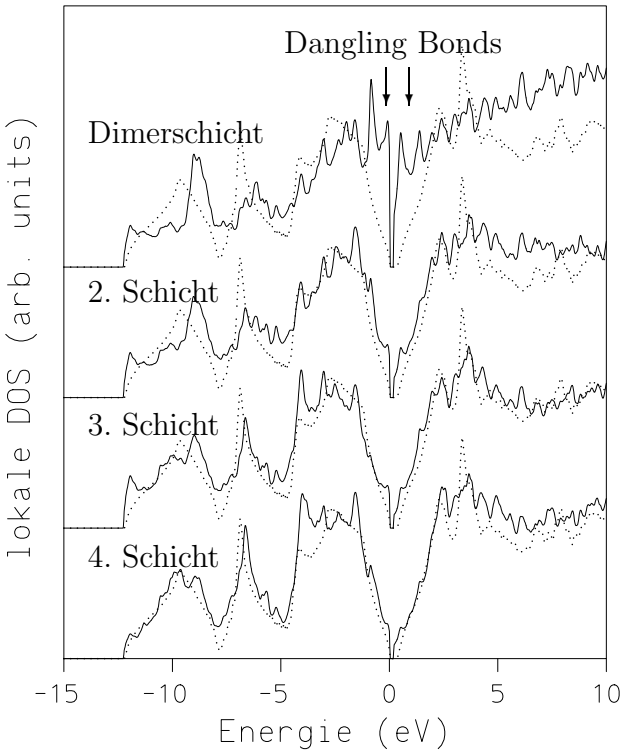


Abbildung 6.4: *Schichtaufgelöste LDA-Zustandsdichte des Si(001)-(2×1) Slabsystems im Vergleich mit der Zustandsdichte des Si Volumenkristalls (gepunktet). Die Oszillationen in der Zustandsdichte des Slabsystems (Periodenlänge ca. 0.5–1 eV) sind auf Interferenzen als Folge der endlichen Slabdicke des Slabs zurückzuführen.*

Der nichtlokale Selbstenergie-Operator besitzt nur eine Reichweite von wenigen Å [God88]. Ein Slab von 10 Å Dicke ermöglicht auch diesem Operator den erforderlichen Übergang von der Oberfläche zum volumenartigen Operator in der Slabmitte.

Die Veränderung der Slabeigenschaften von der Mitte bis zur Oberfläche wird auch bei der energieabhängigen Zustandsdichte deutlich. In Abb. 6.4 wird die lokale LDA-Zustandsdichte des Slabsystems für die Si(001)-(2×1) Oberflächenschicht und die drei darunter liegenden Schichten dargestellt und mit der gesamten Zustandsdichte des Volumenkristalls verglichen. Auch hier ist festzustellen, daß in der Slabmitte das Slabsystem die Volumeneigenschaften gut reproduziert. Je geringer der Abstand zur Oberfläche wird, desto stärker treten in der lokalen Zustandsdichte Unterschiede zum Volumenkristall auf. Neben den volumenartigen elektronischen Zuständen treten zusätzliche Oberflächenzustände in Erscheinung. Bei der hier be-

trachteten Si(001)-(2×1) Oberfläche fallen in der Nähe des fundamentalen Gaps die Zustände der freien, nicht abgesättigten Bindungen (dangling bonds) der Dimeratome besonders markant auf. Ihre Intensität klingt in den Kristall hinein mit einer Halbwertslänge von 2-3 Å ab, so daß sie in der Slabmitte kaum noch festzustellen sind.

Obwohl die lokalen physikalischen Größen wie Ladungsdichte und Potential in der Mitte des Slabs offenbar volumenartig sind, so ist bei der Untersuchung anderer, insbesondere nichtlokaler Volumeneigenschaften die endliche Dicke des Slabsystems zu berücksichtigen. In z-Richtung wird der Slab von Potentialwänden mit ca. 20 eV Höhe eingegrenzt. Dieses Confinement beeinflußt sowohl die volumenartigen als auch die oberflächeninduzierten Zustände des Slabsystems. Wir betrachten zunächst die Auswirkungen auf die volumenartigen elektronischen Zustände. Das Confinement führt dazu, daß in z-Richtung nur diskrete Wellenzahlen zulässig sind. Daher überdecken die volumenartigen Energieeigenwerte des Slabs nicht das gesamte Spektrum des Volumenkristalls. Insbesondere tritt das Volumen-Valenzbandmaximum (VBM), das bei den meisten Materialien am Γ-Punkt liegt, im Spektrum des Slabs nicht auf, denn k_z=0 ist keine zulässige Wellenzahl. Eine einfache Abschätzung ergibt, daß der oberste volumenartige Valenzbandzustand des Slabs um $\hbar^2\pi^2/2m^*L^2$, mit $L = 10$Å also um einige Zehntel eV tiefer liegt als das Volumen-VBM. Das Volumen-VBM ist also im Slabspektrum gar nicht enthalten.

Im realen Oberflächensystem liegt hingegen neben den Oberflächenzuständen das gesamte Volumenspektrum vor. Insbesondere stellt das Volumen-VBM eine natürliche Referenz-Energie für die elektronische Bandstruktur dar. Im Slabsystem liegen nur verfälschte Volumenzustände vor, die keine genaue Aussage über das VBM erlauben. Wir betrachten daher von den elektronischen Zuständen des Slabsystems nur diejenigen, die sich anhand der Lokalisierung ihrer Wellenfunktionen als Oberflächenzustände identifizieren lassen. Die Anknüpfung an das Volumenspektrum erfolgt dann durch Kombination mit der vollständigen, separat berechneten projizierten Bandstruktur des zugrunde liegenden Volumenkristalls. Zwischen den beiden Potentialen des Slabsystems und des Volumenkristalls kann aus rechentechnischen Gründen eine (physikalisch bedeutungslose) Energieverschiebung ΔV gegeneinander auftreten. In der Slabmitte verhält sich das Slab-Potential volumenartig (vgl. Abb. 6.3), und es stimmt in seinem Verlauf mit dem Volumen-Potential überein: $V_{\mathrm{Slab}}(\mathbf{r}) = V_{\mathrm{Bulk}}(\mathbf{r}) + \Delta V$. Durch den Vergleich der beiden Potentiale in diesem Bereich läßt sich die Verschiebung ΔV direkt angeben und ermöglicht eine Angleichung

der Energieskalen der beiden Systeme. In Abb. 6.3 haben wir diese Angleichung bereits vorgenommen.

Während die Verfälschungen der volumenartigen Zustände des Slabs durch das Confinement mit der Einbeziehung der projizierten Volumenbandstruktur gegenstandslos werden, sind die Auswirkungen auf die Oberflächenzustände schwierig zu kontrollieren. Durch die endliche Slabdicke kommt es zu Veränderungen der Oberflächenzustände gegenüber denen der realen Oberfläche. Diese Veränderungen werden um so gravierender ausfallen, je schwächer der jeweilige Zustand an der Oberfläche lokalisiert ist. Zur Untersuchung eines bestimmten Oberflächenzustandes muß der Slab dicker sein als die Ausbreitungstiefe des Zustands in den Kristall hinein. Oberflächenresonanzen, die weit in den Kristall hineinreichen, lassen sich mit einem relativ dünnen Slab prinzipiell nicht korrekt beschreiben. Aber auch Zustände, die deutlich an der Oberfläche lokalisiert sind, erfahren Modifikationen durch die Präsenz der anderen Oberfläche an der Slabunterseite. Wenn die beiden Oberflächen des Slabs gleich aufgebaut sind, so bilden sich beidseitig die gleichen Zustände aus. Durch den Slab hindurch kommt es dann zur Bildung symmetrischer und antisymmetrischer Linearkombinationen. Die Zustände koppeln aneinander, und ihre Energien sind nicht mehr entartet, sondern spalten auf. In einer einfachen Überlegung nehmen wir ein exponentielles Abklingen der Oberflächenzustände in den Slab hinein mit einer Halbwertslänge von etwa 3 Å an. Bei einer Slabdicke von 10 Å ergibt sich ein Überlapp zwischen den Zuständen der beiden Oberflächen von S=0.03, der bei weniger schnellem Abklingen größer ausfällt [Hyb88b]. Die daraus resultierende Kopplung vermag die lokalen Oberflächeneigenschaften, wie etwa die Ladungsdichte, nicht nachhaltig zu stören – sie genügt aber, um eine Energie-Aufspaltung der Oberflächenzustände von bis zu 1 eV zu erreichen. Wir stellen das in Abb. 6.5 dar. Wir vergleichen am Beispiel der Si(001)-(2×1) Oberfläche zwei LDA-Rechnungen miteinander. In der einen Rechnung (gestrichelte Bänder) besitzt der Slab beidseitig die gleiche Oberfläche. Die Bänder spalten dementsprechend auf. Die durchgezogenen Bänder hingegen geben die unverfälschten Energien der ungekoppelten Oberflächenzustände an (wie wir sie berechnen, erläutern wir weiter unten). Der Mittelwert der beiden unbesetzten gekoppelten Bänder liegt um bis zu 0.3 eV über dem zugehörigen unverfälschten Band. Die einfache Mittelwertbildung liefert also ein recht ungenaues Ergebnis für die eigentliche Energie eines ungekoppelten Zustands. Aus den gekoppelten Zuständen läßt sich zwar die Kopplungskonstante S indirekt reproduzieren, die letztlich einen Rückschluß auf die ungekoppelten Energien zuläßt [Hyb88b]. Aber

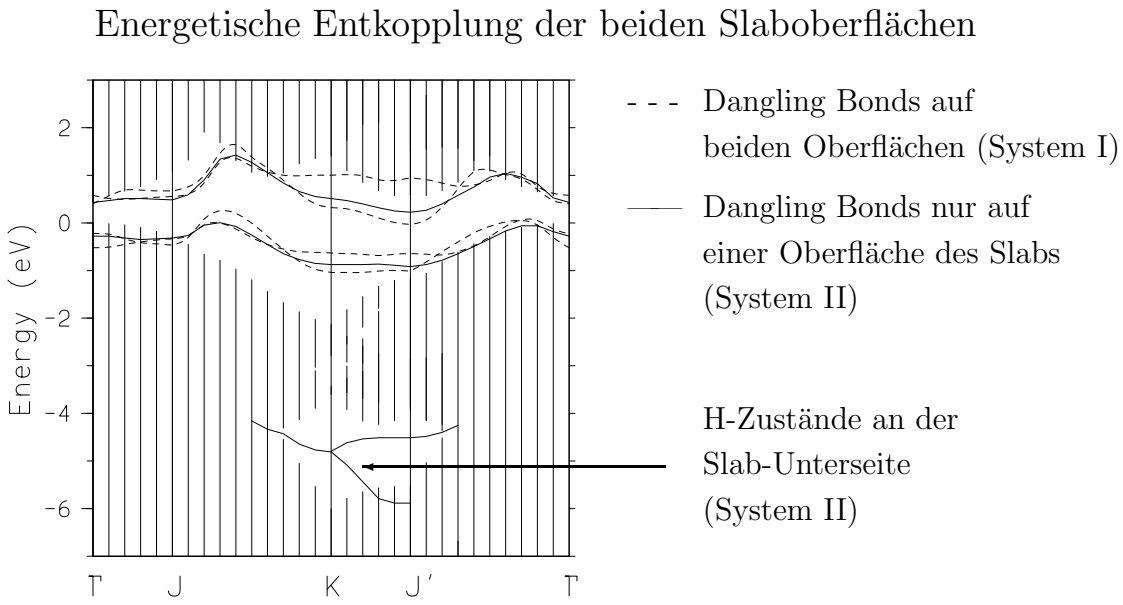


Abbildung 6.5: *Oberflächen-Bandstruktur (LDA) der Si(001)-(2×1) Oberfläche, mit zwei verschiedenen Slabsystemen berechnet. Bei System I befinden sich auf beiden Oberflächen des Slabs Dimere; die zugehörigen Dangling-Bond Zustände koppeln und spalten auf (gestrichelte Bänder). Bei System II wurde die Unterseite des Slabs mit H abgesättigt. Die resultierenden H-Zustände befinden sich bei tieferen Energien (∼ –5 eV) und koppeln nicht an die Dangling Bonds der Oberseite (durchgezogene Kurven).*

auch dieser Zugang bietet nur eine eingeschränkte Genauigkeit.

Wir verfolgen in dieser Arbeit einen anderen Weg zur Ermittlung der ungekoppelten Bandstruktur. Wenn die resonanten Oberflächenzustände auf der Slabunterseite entfernt werden, so wird die Kopplung der Oberflächenzustände der Oberseite mit ihnen vollständig aufgehoben. Die verbleibenden Oberflächenzustände der Oberseite breiten sich dann ungestört aus und klingen in den Slab hinein rasch ab – sofern keine energetische Resonanz mit dem Volumensystem vorliegt. Lediglich durch Reflexion an der Potentialbarriere der Unterseite kommt es zu einer geringen Störung, die aber wesentlich geringer ausfällt als die ursprüngliche resonante Kopplung an die Oberflächenzustände an der Unterseite. Wir erzielen diese Entkopplung, indem die beiden Oberflächen chemisch unterschiedlich konzipiert werden. Wenn etwa eine Oberfläche mit freien Bindungen (Dangling Bonds), deren Energien im Bereich des fundamentalen Gaps liegen, untersucht werden soll, so bietet es sich an, die Unterseite des Slabs mit Wasserstoff abzusättigen. Dadurch werden die dortigen Dangling

Bonds in chemische Bindungen überführt. Die Energien der entsprechenden elektronischen Zustände liegen im Bereich der projizierten Volumenbandstruktur, fern des fundamentalen Gaps (siehe Abb. 6.5). Diese H-Terminierung ermöglicht dann eine ungestörte Untersuchung der Dangling-Bond Zustände auf der oberen Oberfläche.

Kapitel 7

Die Si(001)-(2×1) Oberfläche

Wir untersuchen zunächst reine Halbleiter-Oberflächen ohne Adatome. Dabei stellen wir die Si(001) Oberfläche in den Vordergrund und stellen an ihr die wesentlichen elektronischen Eigenschaften solcher Systeme vor.

Die reine Si(001)-(2×1) Oberfläche ist eine der wichtigsten und am besten untersuchten Halbleiteroberflächen. Sie bietet nicht nur ein ungeheures technologisches Potential, sondern erlaubt auch in besonderem Maß ein Studium des Wechselspiels zwischen der mikroskopischen atomaren Oberflächengeometrie und den elektronischen Eigenschaften. Keine andere Halbleiteroberfläche wurde experimentell und theoretisch so intensiv untersucht wie Si(001). Aus der Fülle der Untersuchungen zeichnet sich folgendes Bild dieser Oberfläche ab.

Ausgehend vom unendlich ausgedehnten Kristall, kann man sich das Entstehen der Oberfläche vorstellen, indem entlang einer (001) Ebene alle Bindungen aufbrechen und eine Hälfte des Systems entfernt wird. An dieser idealen Oberfläche liegen an jedem Oberflächenatom zwei freie Elektronen in Form nicht abgesättigter Bindungen vor. Unabgesättigte Bindungen sind energetisch sehr ungünstig und äußerst reaktiv. Eventuell vorhandene Fremdatome werden an der Oberfläche sofort adsorbiert. An der freien, reinen Oberfläche kann die Zahl der dangling bonds durch Dimerbildung reduziert werden (siehe z. B. [Han87]): je zwei Oberflächenatome nähern sich soweit aneinander an, bis zwischen ihnen eine neue chemische Bindung entsteht. Es kommt zwar dabei an den Atomen der zweiten Schicht zu einer energetisch ungünstigen Verdrehung der Bindungen zur Oberflächenschicht; dies wird aber durch den Energiegewinn der Dimerbindung überkompensiert. Die Dimere sind an der Oberfläche in Reihen angeordnet [Ham86]. Das kann im einfachsten Fall in einer (2×1) Rekonstruktion realisiert werden [Joh90, Joh92, Uhr81].

Eine versetzte Anordnung der Dimere in $c(2\times2)$ Form ist energetisch ungünstiger und wird nicht beobachtet. Eine Vielzahl theoretischer Untersuchungen beziffern den Energiegewinn durch die Dimerisierung auf ca. 2 eV pro Dimer [Krü95, Ram95].

Prinzipiell können die Dimere parallel zur Oberflächenebene orientiert sein (symmetrisches Dimer-Modell, SDM), oder sie können relativ zur Oberflächenebene gekippt sein (asymmetrisches Dimer-Modell, ADM). Im letzteren Fall wird das äußere Dimeratom als Up-Atom und das innere Dimeratom als Down-Atom bezeichnet. Das Kippen der Dimere führt zu grundlegenden Konsequenzen für die elektronischen Zustände. Während die Dangling Bonds im SDM als Folge der Spiegelsymmetrie der Dimere gerade und ungerade Linearkombinationen ausbilden, die als π und π^* Zustände bezeichnet werden, ermöglicht im ADM das Verkippen des Dimers eine räumlichen Separierung der Dangling Bonds in zwei Zustände D_{up} und D_{down}. Der D_{up} Zustand ist überwiegend am Up-Atom lokalisiert, der D_{down} Zustand hingegen am Down-Atom. Der D_{up} Zustand ist dadurch räumlich weiter vom restlichen System entfernt als D_{down}. Er unterliegt einer schwächeren Wechselwirkung mit dem Substrat, und seine Bandstrukturenergie liegt unter der des D_{down} Zustands. Es zeigt sich, daß bei einer Vergrößerung des Kippwinkels der Zustand D_{up} energetisch abgesenkt wird, während D_{down} angehoben wird. Da D_{up} besetzt und D_{down} unbesetzt ist, führt eine Absenkung von D_{up} letztlich auch zu einer Absenkung der Gesamtenergie durch die Verkippung des Dimers [Cha79]. Dadurch erweist sich das ADM als energetisch günstiger als das SDM und wird in den meisten Untersuchungen als reale Oberflächengeometrie für Si(001) favorisiert [Krü95].

Im Rahmen von DFT-LDA Rechnungen ergibt sich der Kippwinkel zu etwa 19° [Krü93, Krü95, Ram95]. Der mit der Kippung verbundene Energiegewinn beträgt 0.15 eV pro Dimer. Die Dimerlänge beträgt etwa 2.25 Å und ist somit etwas kleiner als die Bulkbindungslänge von 2.35 Å. Gegenüber dem idealen Oberflächen-Atomabstand von 3.8 Å kommt es also zu einer erheblichen Verschiebung der Atome. PW-LDA Rechnungen mit geringer Abschneideenergie ergeben deutlich kleinere Kippwinkel von 9–15° [Dab92, Zhu89b]. Daraus wird deutlich, daß das Kippen des Dimers eine Folge der Lokalisierung des Zustands D_{up} ist. Ein PW-Basissatz ohne die stark oszillierenden höheren Wellen weist eine geringere Lokalisierungs-Flexibilität auf und bevorzugt einen geringeren Kippwinkel. Erst bei einer Abschneideenergie von 20 Ryd ergibt sich ein konvergentes Ergebnis für den Kippwinkel [Ram95].

Die Vorstellung asymmetrischer Dimere wird von einer Vielzahl von Experimenten unterstützt. Core-Niveau Spektroskopie-Untersuchungen [Him80] weisen zwei

gegenüber der Volumenlinie verschobene Energieniveaus des Si $2p$ Zustands auf. Das deutet darauf hin, daß an der Oberfläche Si-Atome mit zwei verschiedenen Elektronenkonfigurationen zu finden sind. Der mit einem Kippen verbundene elektronische Ladungstransfer vom Down-Atom zum Up-Atom innerhalb eines Dimers vermag diesen Befund zu erklären. Die Ergebnisse von Streuexperimenten, wie etwa niederenergetische Elektronenstreuung (LEED) [Hol84] und Ionen-Streuung [Tro83], lassen sich mit asymmetrischen Dimeren weit besser vereinbaren als mit symmetrischen Dimeren. Das ADM erlitt einen gewissen Rückschlag durch Rastertunnel-Mikroskopie (STM) Untersuchungen bei Raumtemperatur (RT), die ein symmetrisches Bild der Dimere ergaben [Ham86]. Allerdings steht dieser Befund nicht im Widerspruch zur Vorstellung eines prinzipiell asymmetrischen Dimers, denn dieser vermag bei RT so schnell zwischen seinen beiden möglichen Kipprichtungen zu oszillieren [Dab92, Shk95], daß ein STM Experiment wegen der weit geringeren Zeitauflösung nichts anderes als ein symmetrisches Bild erzielen kann (siehe unten). Ferner kann eine Beeinflussung der Oberflächengeometrie durch die Wechselwirkung zwischen Spitze und Probe während des Experiments nicht ausgeschlossen werden [Cho93]. Neuere experimentelle Untersuchungen bei tieferer Temperatur belegen eindeutig, daß die Dimere gekippt sind [Lan92, Fon93, Mun95, Bul95].

Die Asymmetrie der Dimere führt zu zwei wichtigen Konsequenzen.
(i) Zum einen ergeben die meisten theoretischen Arbeiten [Cha79, Dab92, Krü95], daß mit dem SDM eine metallische Oberflächenbandstruktur verbunden ist, in der die Dangling-Bond Bänder π und π^* überlappen. Das steht in klarem Widerspruch zum experimentellen Befund aus ARPES Untersuchungen [Him79] und Photostrom-Messungen [Mön81], daß die Oberfläche halbleitend ist. In diesem Zusammenhang ist oftmals die Frage aufgeworfen worden, ob die Metallizität auch das Ergebnis einer GW Quasiteilchen-Rechnung wäre, oder ob durch die Quasiteilchen-Korrekturen zur LDA eine energetische Separierung der Bänder erzielt wird, so daß die Quasiteilchen-Bandstruktur des SDM tatsächlich halbleitend würde. Wir haben eine solche GW-Untersuchung durchgeführt. Bevor wir in Abschnitt 7.4 näher darauf eingehen, nehmen wir hier das Ergebnis vorweg: Auch die von uns ermittelte GW-Quasiteilchen-Bandstruktur ist metallisch. Wir verstehen das als ein weiteres starkes Argument gegen das SDM als realistische Oberflächengeometrie.
(ii) Zum anderen ist im Rahmen des SDM keine höhere Rekonstruktion als (2×1) erklärbar. Beugungsexperimente bei Temperaturen unter 200 K zeigen jedoch Reflexe, die über eine (2×1) Rekonstruktion hinausgehen und statt dessen einer $c(4\times2)$ oder

$p(2\times2)$ Rekonstruktion zuzuordnen sind [Pop78, Tab87]. Diese höheren Rekonstruktionen kommen dadurch zustande, daß durch eine entgegengesetzte Orientierung der Kipprichtung benachbarter Dimere (einer antiferromagnetischen Ordnung vergleichbar) die totale Energie reduziert wird. Dieser Energiegewinn beträgt bezüglich der Nachbarschaft innerhalb einer Dimerreihe etwa 0.05 eV pro Dimer und bezüglich der Nachbarschaft zwischen zwei Dimerreihen etwa 0.003 eV pro Dimer [Ram95]. Die Gesamtenergie wird somit minimal, wenn sowohl in Dimerreihen-Richtung als auch senkrecht dazu die Kipprichtung der Dimere alterniert. Die daraus resultierende $c(4\times2)$ Rekonstruktion wird bei tiefen Temperaturen beobachtet [Pop78, Tab87]. Allerdings ist der Energiegewinn durch die höhere Rekonstruktion so gering, daß die ideale $c(4\times2)$ Fernordnung bei höheren Temperaturen verlorengeht. In LEED Experimenten ist diese Fernordnung am $(\frac{1}{4},\frac{1}{2})$ Signal und dazu äquivalenten Signalen feststellbar. Tabata et al. haben gezeigt, daß diese Signale ab einer Temperatur von 150 K immer schwächer werden und bei Raumtemperatur (RT) verschwunden sind [Tab87] . Es handelt sich hierbei um einen Ordnungs-Unordnungs-Phasenübergang zweiter Ordnung. Die Dimerorientierungen an der Oberfläche lassen sich auf ein zweidimensionales Spin-1/2 Isingmodell abbilden. Eine Untersuchung der thermodynamischen Eigenschaften dieses Systems mit Hilfe der Renormierungsgruppentheorie beziffert die Übergangstemperatur des Phasenübergangs auf etwa 250 K [Ihm83]. Bezüglich der Fernordnung ist bei RT nur noch eine nominale (2×1) Rekonstruktion vorhanden (=Dimere in Reihen) [Joh90, Joh92, Uhr81]. Allerdings wird zwischen den Dimeren innerhalb einer Reihe auch bei RT noch eine kurzreichweitige Korrelation vorherrschen [Shk95], die eine entgegengesetzte Kipporientierung benachbarter Dimere favorisiert. Diese Korrelation geht erst bei sehr hohen Temperaturen verloren. Wir werden bei der Diskussion einiger Aspekte der Oberflächenbandstrukturen von Si(001)-(2×1) und Ge(001)-(2×1) darauf zurüchkommen.

Was die lokale Struktur eines einzelnen Dimers betrifft, so kann man auf der Basis des deutlichen Energiegewinns von 0.15 eV durch das Kippen (vergleichbar einer Temperatur von 1800 K) davon ausgehen, daß auch bei RT die Asymmetrie der Dimere erhalten bleibt. Es kann jedoch nicht ausgeschlossen werden, daß der Kippwinkel bei RT anders ist als bei T=0. Eine Studie der Dynamik der Oberfläche zeigt, daß es zu einem dynamischen Flippen eines Dimers zwischen seinen beiden möglichen Kipporientierungen kommen kann [Dab92, Shk95]. Dieses Flippen findet auf einer Zeitskala der Größenordnung 10^{-10}–10^{-8} sec statt. Ein experimentelles Gerät, das nicht über eine derartig hohe Zeitauflösung verfügt, wird automatisch

Geometrie und Brillouinzone der Si(001)-(2×1) Oberfläche

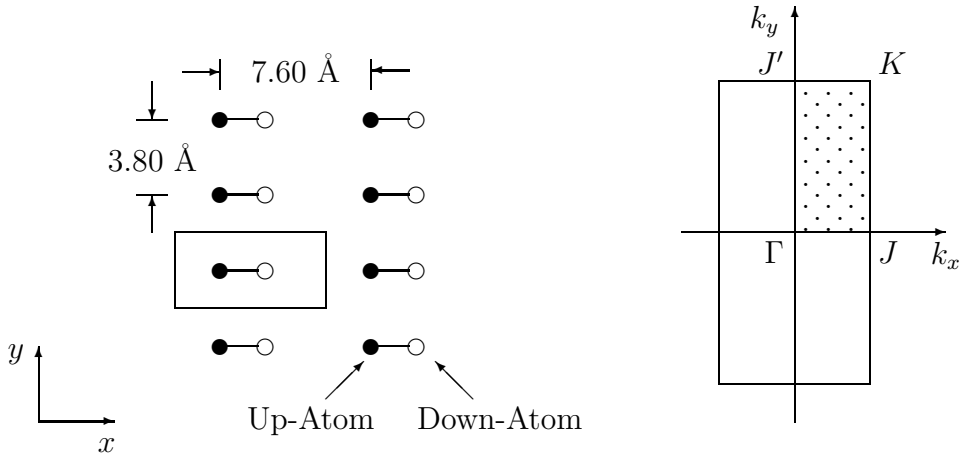


Abbildung 7.1: *Links: Positionen der Oberflächenatome auf der Si(001)-(2×1) Oberfläche im Rahmen des ADM. Die Dimere sind in Reihen angeordnet; alle Dimere besitzen die gleiche Kipp-Orientierung. Rechts: die zugehörige Oberflächen-Brillouinzone. Der irreduzible Teil ist schraffiert.*

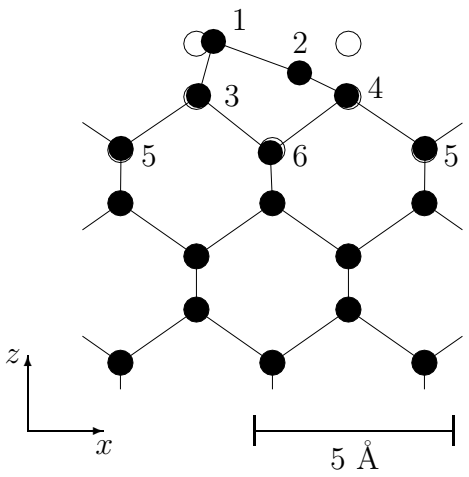
ein symmetrisches Bild ergeben, das einer Mittelung zwischen den zwei Kipporientierungen entspricht. Auf diese Weise wird verständlich, daß STM Aufnahmen bei RT in der Regel Dimere abbilden, die eine symmetrische Form zu besitzen scheinen. Inwieweit die Dynamik der Oberfläche bei RT die Oberflächenbandstruktur von Si(001) beeinflußt, können wir hier nicht untersuchen.

Wir konzentrieren uns im Folgenden auf die grundlegenden elektronischen Eigenschaften der mit asymmetrischen Dimeren bedeckten Si(001) Oberfläche. Zu diesem Zweck gehen wir von einer (2×1) Ordnung aus.

7.1 Das asymmetrische Dimer-Modell (ADM)

Die Si(001)-(2×1) Oberfläche wurde intensiv im Rahmen der LDA untersucht [Büs94, Krü95, Ram95]. In dieser Arbeit konzentrieren wir uns auf die resultierende Bandstruktur und bestimmen die Quasiteilchen-Korrekturen dazu. Dazu folgen wir der in Kap. 2 vorgestellten Vorgehensweise. Die Simulation der Oberfläche geschieht in einem Slabsystem mit acht Schichten Si in einer (2×1)-Einheitszelle mit der theoretischen Gitterkonstante von 5.37 Å. An der Oberfläche liegt eine (2×1) Rekonstruktion mit asymmetrischen Dimeren in einer von M. Sabisch [Sab94] optimierten

Geometrie der Si(001)-(2×1) Oberfläche



Atom	Δx [Å]	Δz [Å]
1	0.426	0.057
2	–1.221	–0.737
3	0.054	0.031
4	–0.047	0.032
5	0.015	0.041
6	–0.052	–0.067

Abbildung 7.2: *Dimerbildung und Relaxation an der Si(001)-(2×1) Oberfläche [Sab94]. Die offenen Kreise entsprechen den Atompositionen der idealen, nicht relaxierten (1×1) Oberfläche. Relaxationen in y-Richtung erweisen sich in der (2×1) Geometrie als energetisch ungünstig und treten nicht auf.*

Geometrie vor. Abb. 7.1 gibt eine schematische Übersicht über die Oberflächengeometrie. Es werden Relaxationen in den drei obersten Schichten berücksichtigt. Die genauen Atompositionen geben wir in Abb. 7.2 an. Die Dimere besitzen eine Bindungslänge von 2.29 Å und einen Kippwinkel von 20.2°. In der unteren Hälfte des Slabs befinden sich die Atome in ihren idealen Positionen. Die freien Bindungen der Si Atome an der Unterseite sind mit zwei H-Atomen je Oberflächenatom abgesättigt. Der Si-H-Bindungsabstand von 1.5 Å entspricht der Bindungslänge in einem SiH_4 Molekül. Auf eine Geometrie-Optimierung an dieser unteren Oberfläche wurde bewußt verzichtet, da sie die Eigenschaften der im Vordergrund stehenden oberen Oberfläche nicht nachhaltig beeinflußt.

In der LDA-Rechnung kommen normerhaltende ab-initio Pseudopotentiale in separabler Form zur Anwendung [Bac82, Kle82]. Für die H-Atome benutzen wir das von Gygi [Gyg93] angegebene Pseudopotential. Pro Atom werden 20 Gauß-Orbitale als LDA-Basisfunktionen verwendet. Die Abklingkonstanten betragen 0.15 und 0.5 für die Si Atome sowie 0.25 und 1.5 für die H-Atome. Die Brillouinzonen-Integration erfolgt mit zwei speziellen Punkten aus dem irreduziblen Teil der Oberflächen-Brillouinzone (SBZ), entsprechend acht Punkten aus der gesamten SBZ. In Abb.

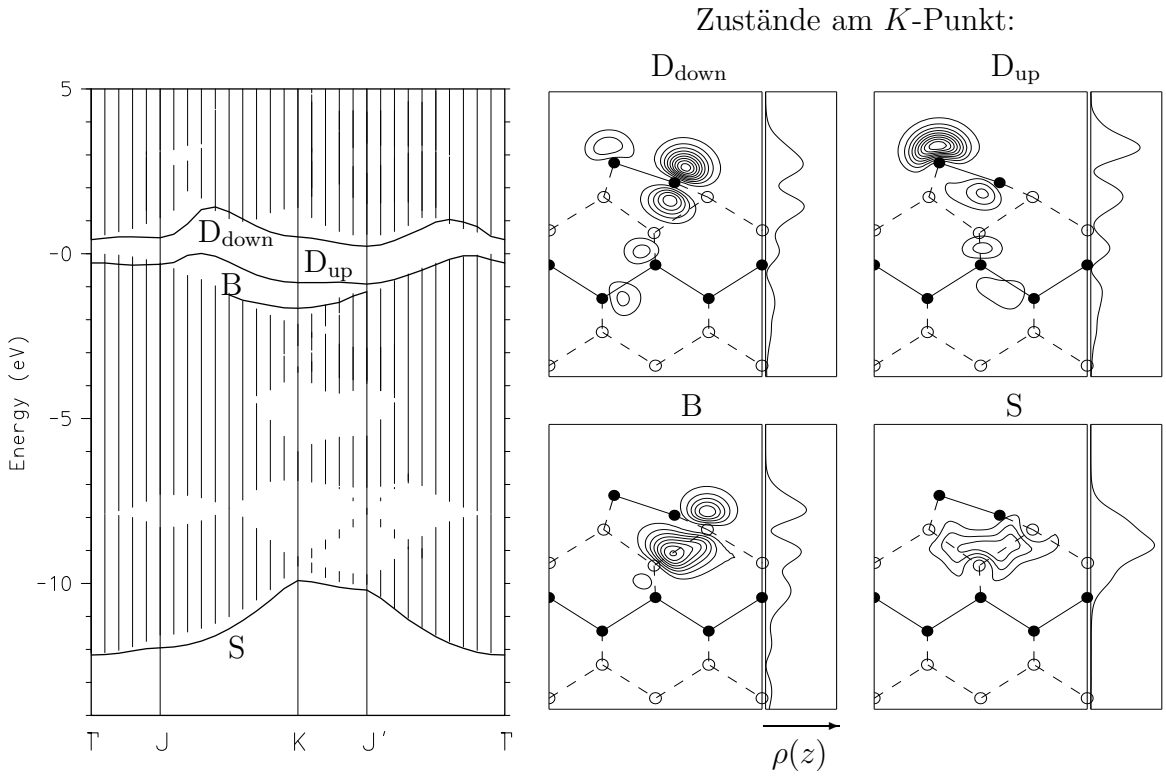


Abbildung 7.3: *Ergebnisse der LDA-Rechnung für Si(001)-(2×1). Die Bandstruktur (links) beschränkt sich auf einige wenige ausgewählte Zustände, die aufgrund ihrer Lokalisierung eindeutig der Oberfläche zugeordnet werden können. Rechts sind für die vier Zustände D_{down}, D_{up}, B und S (jeweils am K-Punkt) die Ladungsdichten dargestellt. Die Zeichenebene verläuft für die Dangling-Bond Zustände durch die Dimeratome, für B und S verläuft sie (parallel verschoben) durch die Si Atome der zweiten Schicht. Anhand der bzgl. x und y gemittelten, z-abhängigen Dichteverteilung $\rho(z)$ eines jeden Zustands (rechts neben jeder Ladungsdichte) ist die Lokalisierung an der Oberfläche ebenfalls deutlich zu erkennen.*

7.3 stellen wir die LDA-Oberflächenbandstruktur dieses Systems in Kombination mit der zugehörigen projizierten Volumen-Bandstruktur (senkrechte Schraffur) vor. Die Zustände, die an der H-terminierten Unterseite des Slabs lokalisiert sind, sind in dieser Abbildung nicht enthalten. Wir beschränken uns auf vier besonders deutliche Oberflächenzustände im fundamentalen Gap und unterhalb der Valenzbänder. Innerhalb des Valenzbandbereichs treten noch weitere oberflächeninduzierte Zustände auf, deren Analyse hier zu weit gehen würde. Die Ladungsdichten $|\psi_{\mathrm{m}\mathbf{k}}(\mathbf{r})|^2$ der vier Zustände werden ebenfalls in Abb. 7.3 dargestellt. Die beiden Dangling-Bond

Zustände fallen wegen der Lage ihrer Energie im fundamentalen Gap besonders deutlich auf. Der energetisch höhere dieser Zustände ist über weite Bereiche der Brillouinzone am Down-Atom lokalisiert; er wird daher als D_{down} bezeichnet. Unterhalb der dritten Si Schicht ist seine Ladungsdichte praktisch auf Null abgeklungen. Ähnlich verhält es sich mit dem Zustand D_{up}, der hauptsächlich am Up-Atom lokalisiert ist und deshalb etwas weiter in das Vakuum hineinragt. Beide Zustände bestehen hauptsächlich aus den Si $3p_z$ Orbitalen der Dimeratome. Wir erwähnen nur, daß es in der zweiten und dritten Si Schicht zu Ladungsumverteilungen und damit zum Auftreten weiterer, dort lokalisierter Oberflächenzustände kommt. Zwei Beispiele sind der Backbond-Zustand B und der Zustand S.

Wir konzentrieren unser Interesse im Folgenden auf die Dangling-Bond Zustände D_{up} und D_{down}. Das hat hauptsächlich zwei Gründe. Zum einen ragen diese beiden p_z-Zustände am weitesten in das Vakuum hinein. Am K Punkt befindet sich das Ladungsdichte-Maximum des D_{up} Zustands 0.5 Å über und das Maximum des D_{down} Zustands nur 0.2 Å unter der Ebene der Up-Atome, während alle anderen Zustände tiefer im Kristall liegen. Die Dangling-Bond Zustände sind daher in oberflächensensitiven Experimenten, etwa im STM, am deutlichsten sichtbar. Zum anderen liegen die zugehörigen Bänder im fundamentalen Gap der Volumenbandstruktur und sind somit entscheidend für den halbleitenden Charakter und die Gapenergie der Oberfläche verantwortlich. Vom Standpunkt einer theoretischen Untersuchung betrachtet, erlauben sie außerdem wegen ihrer Position im Gap das Studium des Selbstenergie-Operators für Zustände in diesem Energiebereich, in dem in der Volumenrechnung keine Zustände auftreten.

Der besetzte D_{up} Zustand weist eine LDA-Bandbreite von 0.95 eV auf. Er besitzt entlang der ΓJ Linie und entlang KJ' nur eine sehr schwache Dispersion. Das liegt daran, daß die D_{up} Atome in dieser kristallographischen Richtung einen großen Abstand von 7.60 Å aufweisen (siehe Abb. 7.1). In y-Richtung beträgt dieser Abstand nur 3.80 Å; die Wechselwirkung der Dangling Bonds ist in dieser Richtung größer, und die Dispersion von D_{up} ist entlang der Linien JK und $J'\Gamma$ wesentlich stärker. Der unbesetzte D_{down} Zustand weist bezüglich seiner Dispersion analoge Eigenschaften auf. Seine gesamte Bandbreite ist mit 1.25 eV etwas größer als die des D_{up} Zustandes. Die beiden Bänder werden von einer fundamentalen Energielücke von 0.20 eV getrennt; die Oberfläche stellt also ein halbleitendes System dar. Die Energielücke hängt empfindlich von der Geometrie der Dimere ab; sie wächst mit steigendem Kippwinkel an. Rechnungen, die einen kleineren Kippwinkel favorisieren,

ergeben eine kleinere Energielücke [Dab92]. Experimentelle Untersuchungen ergeben Gap-Energien zwischen 0.44 und 0.9 eV. Bevor wir im nächsten Abschnitt auf die einzelnen Daten eingehen, stellen wir hier lediglich fest, daß das LDA-Ergebnis von 0.2 eV für das Oberflächengap kleiner als alle gemessenen Werte ist. Vor dem Hintergrund der allgemeinen Unterschätzung von Volumen-Energielücken durch die LDA (Gap-Problem) war auch nichts anderes zu erwarten. Ein detaillierter Vergleich der berechneten Ergebnisse mit experimentellen Daten macht daher an dieser Stelle keinen Sinn. Zunächst muß eine Berechnung der Quasiteilchen-Bandstruktur erfolgen.

7.2 Die Quasiteilchen-Bandstruktur

Quasiteilchen-Bandstrukturen für die Si(001) Oberfläche wurden von Northrup [Nor93] sowie von Kreß et al. [Kre93] bestimmt. Beide Untersuchungen beruhen auf dielektrischen Modellfunktionen. Wir stellen in dieser Arbeit die unseres Wissens erste GW-Rechnung für Si(001)-(2×1) vor, in der die dielektrische Matrix im Rahmen der RPA berechnet wird. Dies ermöglicht u.a. einen Vergleich mit Bandstruktur-Resultaten, wie sie aus dielektrischen Modellfunktionen folgen. Die Realisierung der GW-Näherung folgt der in Teil I der Arbeit vorgestellten Vorgehensweise. Wir stellen die Zweipunktfunktionen mit Gauß-Orbitalen dar (GO-GWA). Wir verwenden für die Si Atome s-, p-, d- und s^*-Orbitale mit den Abklingkonstanten 0.15 und 0.5 sowie zusätzlich noch f-Orbitale mit der Abklingkonstanten 0.5 (vgl. Abschnitt 3.4). An den H-Atomen der Slabunterseite, die für die Berechnung des Selbstenergie-Operators an der Slaboberseite relativ unbedeutend sind, verwenden wir s-, p-, d- und s^*-Orbitale mit der Abklingkonstanten 0.3. Insgesamt besteht der GO-GWA-Basissatz damit aus 520 Funktionen. Ein PW-GWA Basissatz gleicher Güte müßte mindestens 1000 ebene Wellen enthalten. Daher erfordert unser GO-GWA Zugang insgesamt kürzere Rechenzeiten als die PW-GWA. Im vorliegenden Fall einer (001)-(2×1) Oberfläche wird der Gewinn an Rechenzeit noch durch die Orthogonalität der Gittervektoren verstärkt. Für orthogonale Gittervektoren läßt sich in der Berechnung der Dreizentrenintegrale (3.14) die sechsdimensionale Summe $\sum_{\mathbf{R},\mathbf{R}'}$ in drei zweidimensionale Summen $\sum_{\mathrm{R}_i,\mathrm{R}'_i}$ (i=1,2,3 kartesische Richtungen) separieren. Voraussetzung ist die Verwendung kartesischer Gauß-Orbitale als Basisfunktionen. Diese Berechnung arbeitet um einen Faktor 2–5 schneller als eine nicht separierte Summation. Zur Berechnung der dielektrischen Matrix verwenden wir zwei spezielle Punkte im irreduziblen Teil der Oberflächen-Brillouinzone. Nur zur Bestimmung

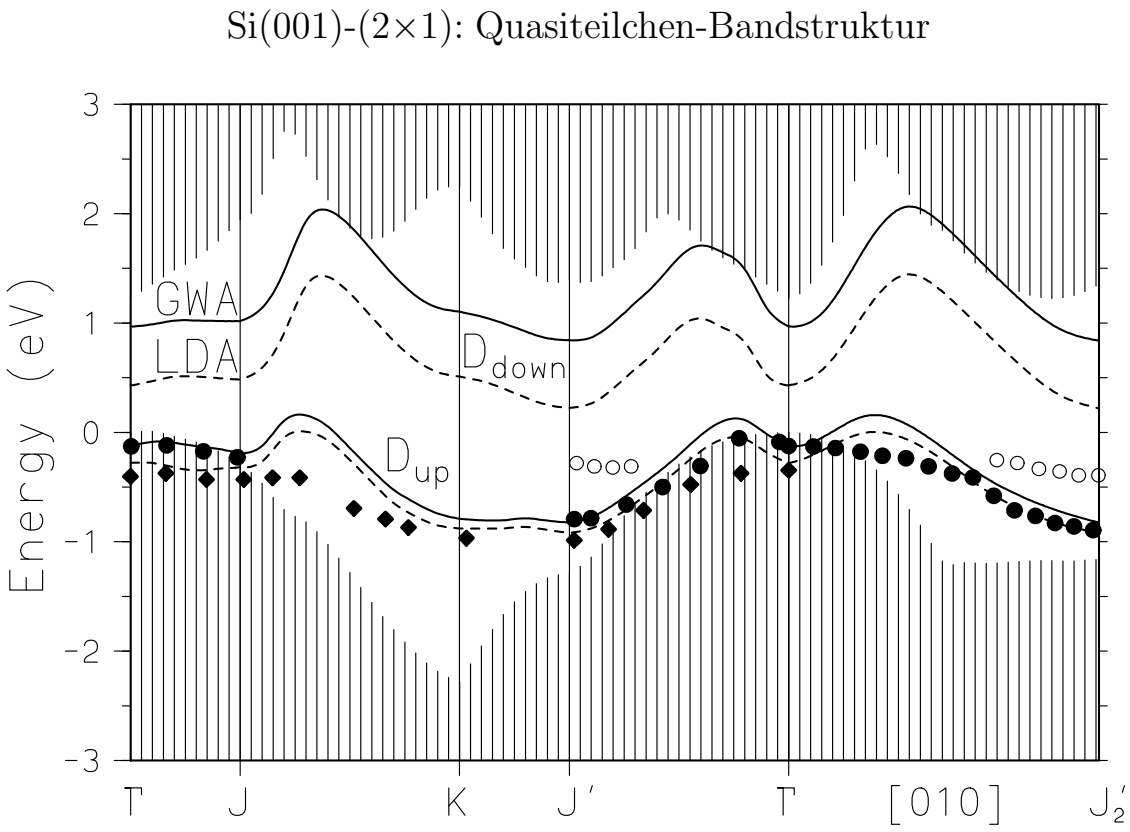


Abbildung 7.4: *Quasiteilchen-Bandstruktur der Dangling-Bond Zustände an der Si(001)-(2×1) Oberfläche. Wir vergleichen unsere GW-Ergebnisse mit den zugrunde liegenden LDA-Energien (gestrichelte Bänder) und mit ARPES-Daten von Uhrberg et al. (Rauten; [Uhr81]) und von Johansson et al. (Kreise; [Joh90]).*

des Kopfes der dielektrischen Matrix im Falle $\mathbf{q}\rightarrow\mathbf{0}$ werden acht spezielle Punkte herangezogen.

Wir bestimmen nach Gl. (2.19) die Quasiteilchen-Korrekturen zur LDA-Bandstruktur der Dangling-Bond Zustände. Die resultierenden Quasiteilchen-Energien der Dangling Bonds stellen wir in Abb. 7.4 dar, zusammen mit den zugrunde liegenden LDA-Energien (gestrichelt) und mit experimentellen Daten der winkelaufgelösten Photoemission (ARPES) [Uhr81, Joh90]. Der D_{up} Zustand wird nominal um –0.35 bis –0.20 eV zu tieferen Energien verschoben. Da gleichzeitig in einer äquivalenten GW-Rechnung für den Volumenkristall das Valenzband-Maximum (VBM) um –0.41 eV zu niedrigeren Energien verschoben wird, ergibt sich insgesamt eine effektive *Anhebung* des D_{up} Zustandes um + 0.05 bis +0.20 eV. Die Bandbreite wird gegenüber der LDA um 0.05 eV auf 1.00 eV erhöht. Insgesamt wird die Dispersion von

D_{up} durch die Quasiteilchen-Korrekturen kaum verändert, und die GWA-Energien von D_{up} unterscheiden sich von den LDA-Resultaten nur marginal. Die berechneten Energien des D_{up} Zustandes stimmen gut mit den experimentellen Daten überein (sowohl in LDA als auch in GWA), besonders mit denen von Johansson et al. (volle Kreise; [Joh90]). Das gilt besonders für die GW-Daten entlang der ΓJ Linie und am J' Punkt. Zwischen J und K, zwischen J' und Γ und entlang der [010] Richtung liegen die Quasiteilchenenergien etwas über den experimentellen Daten. Insbesondere werden die ausgeprägten Maxima der LDA- und GWA-Bandstruktur entlang dieser Richtungen, die auch von anderen Autoren erzielt wurden [Dab92, Krü95], experimentell nicht beobachtet. Die beiden experimentell ermittelten Bandstrukturen aus [Uhr81] und [Joh90] unterscheiden sich um eine nahezu konstante Verschiebung von 0.3 eV. Das könnte darauf zurückzuführen sein, daß die Bestimmung des Abstandes zwischen dem Ferminiveau und dem VBM experimentell schwierig ist. In der Nähe von J' und von J_2' treten im Photoemissionsspektrum von Johansson et al. [Joh90] einige Zustände bei einer Energie von etwa –0.3 eV auf, die nicht zur theoretischen Bandstruktur passen. Sie lassen sich nicht als Zweidomäneneffekt erklären, denn die in dieser Messung vorliegende Si(001) Probe ist eindomänig [Joh90]. Eine mögliche Interpretation fußt auf der entgegengesetzten Kipporientierung benachbarter Dimere (siehe oben). Sie führt zwar bei RT nicht zu einer Fernordnung über die (2×1) Rekonstruktion hinaus (LEED Experimente ergeben nur eine (2×1) Struktur), wohl aber zu einer lokalen Nahordnung in Form einer Korrelation. In einer $p(2\times2)$ Rekonstruktion fallen $\Gamma_{2\times1}$ und $J'_{2\times1}$ beide auf den Punkt $\Gamma_{p(2\times2)}$ und sind somit äquivalent (in Abb. 7.5 durch Pfeile angedeutet), ebenso wie $J_{2\times1}$ und $K_{2\times1}$. Bei einer $c(4\times2)$ Geometrie sind $\Gamma_{2\times1}$ und $K_{2\times1}$ äquivalent, ebenso $J_{2\times1}$ und $J'_{2\times1}$ (siehe Abb. 7.5). Wenn lokal eine höhere Rekonstruktion vorliegt, so können als Folge von Umklapp-Prozessen in der (2×1) Bandstruktur am J' Punkt Zustände erkennbar sein, die ihrer Natur her vom $\Gamma_{2\times1}$ Punkt (bei lokaler $p(2\times2)$ Geometrie) oder vom $J'_{2\times1}$ Punkt (bei lokaler $c(4\times2)$ Geometrie) herrühren. Tatsächlich beträgt die Quasiteilchen-Energie des D_{up} Zustandes in unserer (2×1) Rechnung –0.15 eV am Γ-Punkt bzw. –0.20 eV am J-Punkt. Es sollte erwähnt werden, daß im Rahmen einer höheren Rekonstruktion ($c(4\times2)$ oder $p(2\times2)$) die Bandbreite des D_{up} Bandes reduziert wird [Nor93], weil die Up-Atome dann räumlich weiter voneinander entfernt sind als in der (2×1) Ordnung. Eine lokale höhere Rekonstruktion könnte der Grund sein, warum die experimentell beobachtete Bandbreite des D_{up} Zustandes geringer ist als die von uns in (2×1) berechnete.

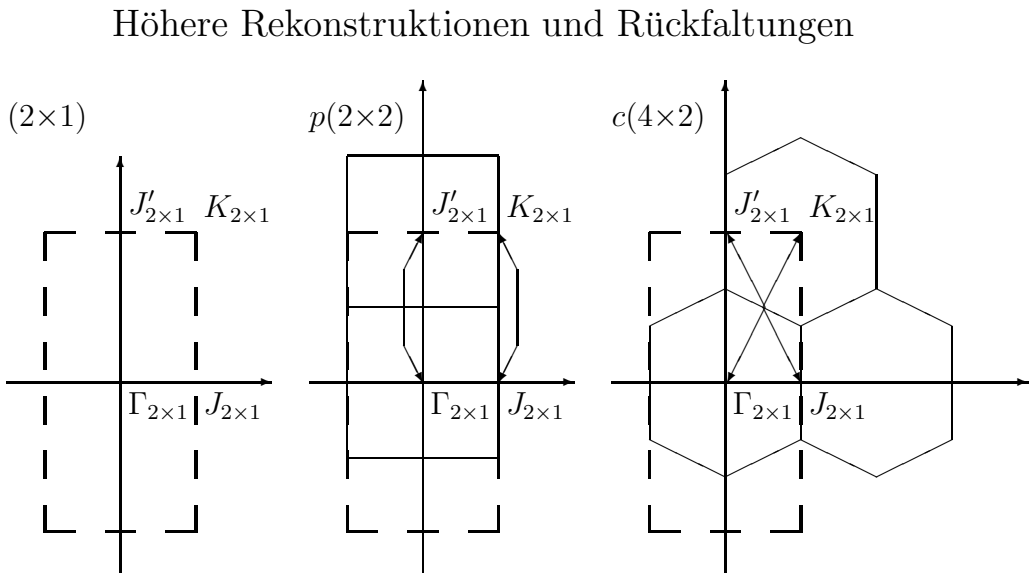


Abbildung 7.5: *Rückfaltung der (2×1) Brillouinzone (gestrichelt) innerhalb der höheren Rekonstruktionen p(2×2) (Mitte; durchgezogene Brillouinzone) und c(2×2) (rechts). Die durch Pfeile verbundenen Symmetriepunkte der (2×1) Geometrie fallen auf die gleichen Punkte der höheren Rekonstruktion und werden dadurch äquivalent.*

Der unbesetzte D_{down} Zustand erfährt durch die Quasiteilchen-Korrekturen eine nominelle Anhebung zwischen +0.10 und +0.25 eV. Bezüglich des VBM führt das zu einer effektiven Anhebung von +0.50 bis +0.65 eV. Die Bandbreite wird gegenüber der LDA ein wenig reduziert; sie beträgt nun 1.20 eV. Es gibt in der Literatur nur wenige inverse Photoemissions-Resultate (ARIPES) für die Si(001) Oberfläche [Joh92]. Da die in [Joh92] angegebene absolute Energieskala bezüglich des VBM unklar ist, beziehen wir diese Daten nicht in die Diskussion mit ein. Die Daten unterstützen allerdings die Vermutung, daß lokale höhere Rekonstruktionen an der Si(001) Oberfläche vorliegen, denn im Spektrum treten *zwei* unbesetzte Bänder deutlich hervor [Joh92].

Für die Quasiteilchen-Bandstruktur ergibt sich ein indirektes fundamentales Gap von 0.70 eV. Dieser Wert ist um 0.50 eV größer als das LDA-Gap (0.20 eV), aber wesentlich geringer als das Volumen-Gap (1.23 eV in unserer GWA-Rechnung). Die Oberflächenzustände reduzieren die Volumen-Energielücke erheblich. Die uns vorliegenden experimentellen Ergebnisse für die Energielücke dieser Oberfläche wurden mit verschiedenen Verfahren ermittelt und streuen über einen weiten Bereich. Die konsistenteste Vergleichbarkeit fundamentaler Gap-Energien mit gemessenen Daten

besteht in Photostrom-Experimenten. Eine Messung von Mönch et al. [Mön81] ergibt eine Gap-Energie von E_{gap}=0.64 eV. Unser Ergebnis stimmt mit diesem Meßwert sehr gut überein. In optischer Absorptionsspektroskopie wird mit 0.44 eV ein wesentlich kleineres Gap gemessen [Cha83]. Die Ursache für die Diskrepanz von 0.2 eV zur Photostrom-Messung und zu unserer Rechnung ist vermutlich darin begründet, daß durch Oberflächenexzitonen die optische Absorptionsschwelle abgesenkt wird [Nor91]. Die im Vergleich zu Volumenexzitonen (Bindungsenergie in Si: 15 meV) weit höhere Bindungsenergie in der Größenordnung von 200 meV erklärt sich aus der Zweidimensionalität der Oberfläche, aus der geringeren dielektrischen Abschirmung und den größeren effektiven Massen an der Oberfläche [Pan95]. Wir erinnern daran, daß die GW-Näherung an sich keine Aussage über Exzitonen treffen kann; sie kann allerdings verläßliche Einteilchen-Daten zur Behandlung von Exzitonen bereitstellen [Str82, Oni95]. Solche Untersuchungen werden von uns hier nicht verfolgt.
Schließlich zeigen Tunnel-Spektroskopie (STS) Untersuchungen eine Energielücke von 0.9 eV [Ham89], die größer ist als das von uns berechnete fundamentale Gap. Der Tunnelstrom in einem solchen Experiment wird im wesentlichen von der lokalen Zustandsdichte $\rho(\mathbf{r}, E)$ der Probe am Ort der Tunnelspitze bestimmt [Ter85]:

$$\rho(\mathbf{r}, E) = \sum_{\mathrm{m}\mathbf{k}_\parallel} |\psi_{\mathrm{m}\mathbf{k}_\parallel}(\mathbf{r})|^2 \delta(E - E_{\mathrm{m}\mathbf{k}_\parallel}) \quad . \tag{7.1}$$

Zu dieser Zustandsdichte tragen die Oberflächenzustände $\mathrm{D_{up}}$ und $\mathrm{D_{down}}$ wegen ihrer Lokalisierung an der Oberfläche am stärksten bei. Außerdem spielt das asymptotische Verhalten der Wellenfunktionen $\psi_{\mathrm{m}\mathbf{k}_\parallel}(\mathbf{r})$ für große Abstände von der Oberfläche eine entscheidende Rolle [Ter85]. Der Ansatz

$$\psi_{\mathrm{m}\mathbf{k}_\parallel}(\mathbf{r}) \sim e^{i\mathbf{k}_\parallel \mathbf{r}_\parallel} e^{-\kappa z} \tag{7.2}$$

führt für die exponentielle Abklingkonstante κ auf die Beziehung

$$\frac{\hbar^2}{2m}\kappa^2 = \frac{\hbar^2}{2m}\mathbf{k}_\parallel^2 + W - E_{\mathrm{m}\mathbf{k}_\parallel} \quad , \tag{7.3}$$

wobei W die Ionisierungsenergie der Oberfläche bedeutet. Die Dangling-Bond Zustände weisen nur eine schwache Dispersion $E_{\mathrm{m}\mathbf{k}_\parallel}$ auf. Das führt dazu, daß ihre Wellenfunktionen am Γ-Punkt wegen des kleinsten Wellenvektors $\mathbf{k}_\parallel$ die kleinste Abklingkonstante κ aufweisen und am weitesten in das Vakuum hineinreichen. Das Energiegap, das im Tunnelexperiment festgestellt wird, wird also von den Bandstrukturenergien der beiden Oberflächenzustände am Γ-Punkt dominiert. Am Γ-Punkt

erhalten wir in unserer Rechnung ein direktes Gap von 1.10 eV zwischen den D_{up} und D_{down} Zuständen, in guter Übereinstimmung mit dem experimentellen Tunnelgap von 0.9 eV [Ham89].

Das direkte Gap zwischen D_{up} und D_{down} reicht von 1.10 eV (an Γ) bis 2.00 eV (entlang JK); Die Mittelung über die Oberflächen-Brillouinzone ergibt einen Wert von 1.6 eV. Dieses Resultat stimmt gut mit dem experimentellen Befund von 1.6–1.7 eV für die mittlere Energie des direkten optischen Übergangs überein [Cha83].

Im Folgenden sollen die Selbstenergien der Dangling-Bond Zustände näher untersucht werden. Wir greifen hierzu die Diskussion aus Kap. 5 wieder auf (vgl. Abb. 5.1). und beziehen die hier erzielten Ergebnisse für die Oberflächenzustände mit ein. Wir stellen in Abb. 7.6 die Austausch-Energie, die Korrelationsenergie und die gesamte Selbstenergie einzelner Zustände in Abhängigkeit von der jeweiligen LDA-Energie dar. Dabei beziehen sich Punkte auf Si Volumenzustände und Kreuze auf die Dangling-Bond Zustände der Si(001)-(2×1) Oberfläche.

Die Austausch-Energie (a) des besetzten D_{up} Zustandes ist schwächer als die der besetzten Volumenzustände; sie liegt über der des VBM. Die Korrelationsenergie (b) ist kleiner als die der besetzten Volumenzustände. Insgesamt reicht die Selbstenergie (c) von D_{up} von –11.4 bis –10.2 eV und fällt somit geringer aus als die der entsprechenden Volumenzustände (–11.65 eV am VBM). Die Austauschenergie des unbesetzten D_{down} Zustandes ist stärker als die der unbesetzten Volumenzustände; sie liegt unterhalb der Austauschenergie des Leitungsband-Minimums (CBM). Die Korrelationsenergie fällt hingegen etwas schwächer aus und liegt über der des CBM. Insgesamt ergibt sich eine Selbstenergie zwischen –9.9 eV und –9.2 eV für D_{down}, die etwas größer ausfällt als für das CBM (–9.05 eV).

In der Volumenrechnung differieren die Selbstenergien des VBM und des CBM um 2.60 eV. Dieser Unterschied stellt den wesentlichen Beitrag zur Existenz einer fundamentalen Energielücke in der Volumen-Bandstruktur dar. Dieser große Unterschied zwischen den Selbstenergien der besetzten und unbesetzten Bänder wird an der Oberfläche stark reduziert. Dies führt dazu, daß in der Oberflächen-Bandstruktur die Energielücke zwischen den Dangling Bonds geringer ausfällt als die Energielücke zwischen den Volumenzuständen. Das betrifft sowohl die minimalen Gaps (0.70 eV an der Oberfläche gegenüber 1.23 eV im Volumen) als auch die direkte Energielücke (im Mittel 1.6 eV gegenüber 3.5 eV).

Darüber hinaus diskutieren wir in Abb. 7.7 die Quasiteilchen-Korrekturen der Zustände in Abhängigkeit von ihren LDA-Energien. Die Quasiteilchen-Korrekturen

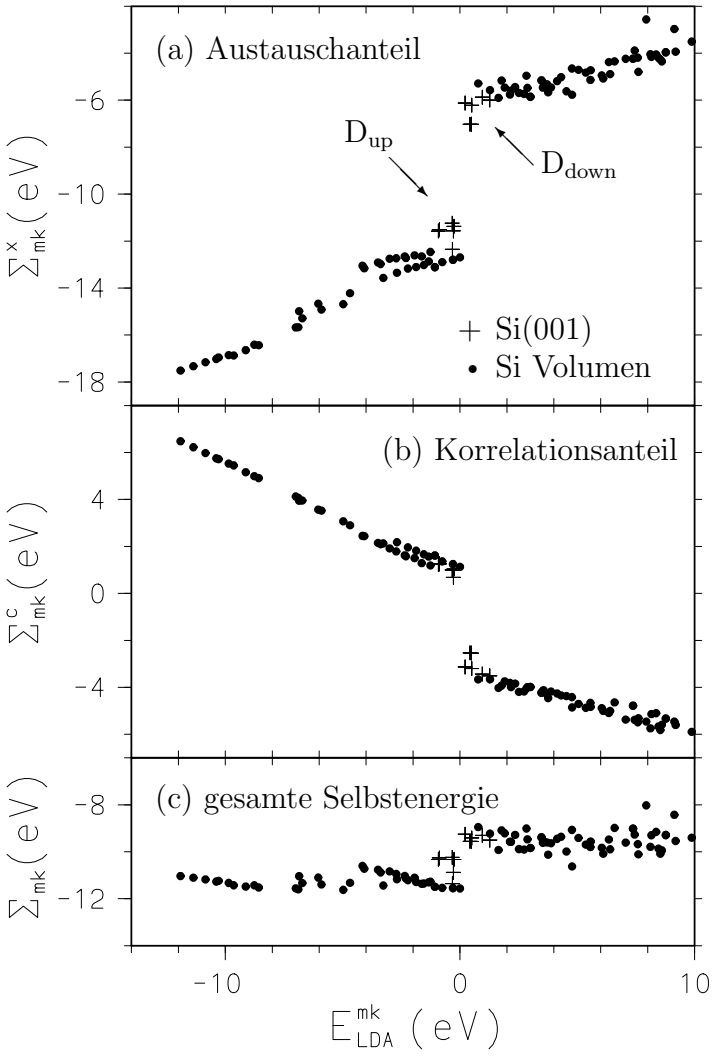


Abbildung 7.6: *Selbstenergien der Dangling-Bond Zustände an der Si(001)-(2×1) Oberfläche (Kreuze) im Vergleich zu den entsprechenden Daten im Si Volumenkristall (Punkte). Der Austauschanteil (a), der Korrelationsanteil (b) und die gesamte Selbstenergie (c) werden gegen die LDA-Energie des jeweiligen Zustandes aufgetragen.*

für die Dangling-Bond Zustände sind etwas geringer als für die Volumenzustände. Die Korrektur-„Stufe“ aus Abb. 3.3 wird dadurch reduziert, und das Oberflächen-Gap erfährt mit +0.50 eV eine geringere Aufweitung als das Volumengap (+0.77 eV). Auch in dieser Abbildung passen die Oberflächendaten gut zu den Volumendaten. Die deutliche Stufenfunktion zwischen den besetzten und unbesetzten Volu-

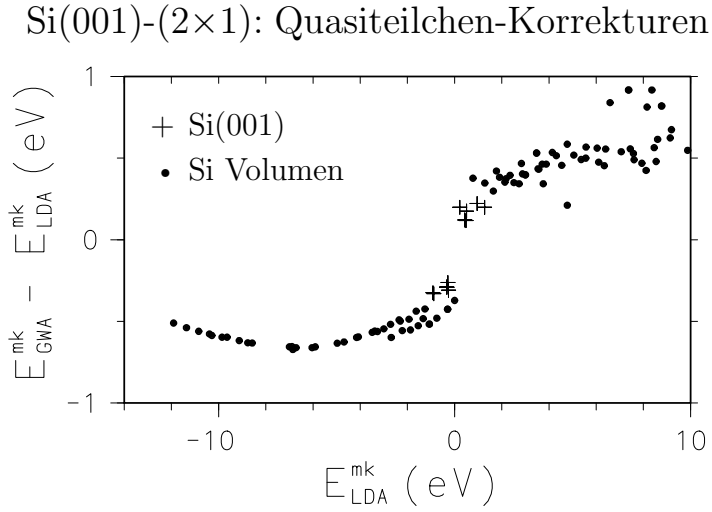


Abbildung 7.7: *Quasiteilchen-Korrekturen* $E^{\mathrm{GWA}}_{\mathrm{m}\mathbf{k}} - E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} = \langle \mathrm{m}\mathbf{k}|\Sigma(E^{\mathrm{GWA}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}}|\mathrm{m}\mathbf{k}\rangle$ *für die Dangling-Bond Zustände an der Si(001)-(2×1) Oberfläche (Kreuze) im Vergleich zu den entsprechenden Si Volumendaten (Punkte). Die Korrekturen sind gegen die LDA-Energie des jeweiligen Zustandes aufgetragen.*

menzuständen wird durch die Korrekturen der Dangling Bonds, deren Bandstruktur-Energien im Gap liegen, geglättet. Auch die Renormierungs-Konstanten der Oberflächen- und der Volumenzustände unterscheiden sich nur sehr wenig. Für die Dangling-Bond Zustände ergeben sich Werte zwischen 0.76 und 0.78. Für die Volumenzustände nahe des Gaps finden wir etwas höhere Werte von 0.79 bis 0.81.

Der gesamte Vergleich der Selbstenergien der Oberflächen- und der Volumenzustände muß vor dem Hintergrund beurteilt werden, daß die Dangling-Bond Zustände deutlich an der Oberfläche lokalisiert sind. Die Abschirm-Eigenschaften an der Oberfläche unterscheiden sich von denen im Volumen. Dementsprechend ist der Selbstenergie-Operator $\Sigma(\mathbf{r}, \mathbf{r}', E)$ an der Oberfläche anders strukturiert als im Volumenkristall. Aus diesem räumlichen Unterschied und aus den verschiedenen Energien der Oberflächen- und der Volumenzustände erklären sich die Unterschiede zwischen den Selbstenergien $\Sigma_{\mathrm{m}\mathbf{k}}$ der einzelnen Zustände und zwischen ihren Korrekturen. Vor diesem Hintergrund ist es ein erstaunliches Resultat, daß die Daten der Dangling-Bond Zustände derartig gut zu denen der Volumenzustände passen. Daran zeigt sich, daß die Oberflächenzustände keine völlig eigenständigen Phänomene darstellen, sondern sich in ihren Eigenschaften eng an dem zugrunde liegenden Substrat orientieren – zumindest bei der hier betrachteten reinen Oberfläche. Insbesondere

setzen sich die Oberflächenzustände aus den gleichen Si Atomorbitalen zusammen wie die Volumenzustände. Ferner untermauern unsere Ergebnisse die Tragfähigkeit unseres Zugangs zur Bestimmung von Quasiteilchen-Eigenschaften an Oberflächensystemen.

7.3 Dielektrische Modellfunktionen

In diesem Abschnitt untersuchen wir anhand der Si(001)-(2×1) Oberfläche, inwieweit sich dielektrische Modellfunktionen zur Berechnung von Quasiteilchen-Energien von Oberflächenzuständen einsetzen lassen. Der wesentliche Unterschied zwischen einer dielektrischen Modellfunktion und dem aufwendigen RPA-Ausdruck (2.8) besteht darin, daß die elektronische Struktur des Systems in das Modell in stark vereinfachter Weise eingeht, manchmal nur in Form der elektronischen Ladungsdichte. Dabei geht die Zusammensetzung der Ladungsdichte aus den einzelnen besetzten Zuständen ebenso verloren wie jede Information über die unbesetzten Zustände und das Energiespektrum. Nur die makroskopische dielektrische Konstante, die meistens berücksichtigt wird, enthält eine gewisse Information über die Bandstruktur und das fundamentale Gap. Die meisten Modellfunktionen erlauben zunächst nur eine Bestimmung der statischen Abschirmeffekte bei der Frequenz Null. Eine Erweiterung zur dynamischen dielektrischen Matrix kann dann nur im Rahmen des einfachsten Plasmon-Pol Modells (I) oder (II) erfolgen.

Die Bestimmung der dielektrischen Matrix im Rahmen der RPA zeichnet sich besonders dadurch aus, daß sie auf alle Elektronensysteme gleichermaßen anwendbar ist. Dies bezieht sich auch auf Situationen, in denen die normale kovalente Bindungsstruktur des Volumenhalbleiters nicht mehr vorliegt, wie etwa bei Defekten, Grenzflächen und Oberflächen. Ob eine dielektrische Modellfunktion in solchen Systemen zu quantitativ richtigen Aussagen führen kann, muß sich erst im Vergleich mit der tatsächlich berechneten dielektrischen Matrix herausstellen.

Die vereinfachte Struktur einer dielektrischen Modellfunktion erlaubt eine wesentlich schnellere Bestimmung der Abschirmeffekte, als dies mit einer RPA Rechnung möglich ist. Hier liegt das enorme Potential der Modellfunktionen bei der Anwendung auf große Systeme, bei denen die Bestimmung der dielektrischen Matrix im Rahmen der RPA aus Rechenzeitgründen ausscheidet. Wir haben in früheren Arbeiten verschiedene dielektrische Modellfunktionen zur Berechnung der Quasiteilchen-Bandstruktur von Volumenhalbleitern untersucht [Roh93a, Roh93b]. Es stellte sich

heraus, daß insbesondere die beiden in [Roh93a, Roh93b] als B2 und C2 bezeichneten Modellfunktionen zu Quasiteilchen-Energien führten, die gut mit den Ergebnissen übereinstimmten, die mit Hilfe der vollen RPA-Funktion berechnet wurden. Wir wollen hier nun untersuchen, ob diese Modellfunktionen auch zur Berechnung der Quasiteilchen-Korrekturen für die Dangling-Bond Zustände an der Si(001) Oberfläche geeignet sind. Falls sich das als richtig herausstellt, so wäre damit der Anwendbarkeitsbereich dieser Funktionen erheblich erweitert. Wir stellen hier die beiden Modellfunktionen nur kurz vor. Eine ausführliche Diskussion findet sich in [Roh93a].

(**B2**) Hybertsen und Louie schlagen ein nichtdiagonales dielektrisches Modell der folgenden Art vor [Hyb88a]:

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}') = \frac{1}{2}\epsilon_{\mathrm{LL}}^{-1}(|\mathbf{r}-\mathbf{r}'|,\rho(\mathbf{r})) + \frac{1}{2}\epsilon_{\mathrm{LL}}^{-1}(|\mathbf{r}-\mathbf{r}'|,\rho(\mathbf{r}')) \quad . \tag{7.4}$$

Dieses Modell berücksichtigt in seiner Struktur die Inhomogenität der Ladungsdichte $\rho(\mathbf{r})$. Es basiert auf der diagonalen Modellfunktion $\epsilon_{\mathrm{LL}}^{-1}(|\mathbf{r}-\mathbf{r}'|,\rho(\mathbf{r}))$ von Levine und Louie [Lev82], die für ein „halbleitendes homogenes Elektronengas“ aufgestellt wurde. Diese diagonale Funktion fußt auf der RPA, bezieht jedoch das fundamentale Gap mit ein, das den Halbleiter vom Elektronengas unterscheidet. Für ϵ_{LL} ergibt sich der Ausdruck

$$\begin{aligned}\epsilon_{\mathrm{LL}}(q,\bar{\rho}) &= 1 + \frac{2}{\pi q_{\mathrm{F}}}\left[\frac{1}{Q^2} - \frac{\lambda}{2Q^3}\left(\arctan\frac{2Q+Q^2}{\lambda} + \arctan\frac{2Q-Q^2}{\lambda}\right)\right. \\ &\quad \left. + \left(\frac{\lambda^2}{8Q^5} + \frac{1}{2Q^3} - \frac{1}{8Q}\right)\ln\left(\frac{\lambda^2+(2Q+Q^2)^2}{\lambda^2+(2Q-Q^2)^2}\right)\right] \quad . \end{aligned} \tag{7.5}$$

Hierbei ist $\bar{\rho}$ die mittlere Elektronendichte und q_{F} die entsprechende Fermi-Wellenzahl. Für den Parameter Q setzt man $Q{=}q/q_{\mathrm{F}}$ ein. Der Parameter λ, der formal die Information über das Gap enthält, wird an die dielektrische Konstante des Systems angepaßt; man wählt $\lambda{=}\frac{\omega_{\mathrm{Pl}}}{\omega_{\mathrm{F}}}\sqrt{\epsilon_\infty - 1}$, wobei ω_{Pl} die Plasmafrequenz und ω_{F} die Fermifrequenz ist.

(**C2**) Ein anderes Modell geht auf Vorschläge von Bechstedt et al. [Bec81] sowie von Falter et al. [Fal91, Fal92] zurück. Anders als das Modell (B2) wird es direkt im Fourierraum konzipiert. Die Diagonalelemente der PW-Darstellung besitzen folgende Form:

$$\epsilon_{\mathbf{G},\mathbf{G}'}(q,0) = \delta_{\mathbf{G},\mathbf{G}'} + \frac{\omega_{\mathrm{Pl}}^2}{\tilde{\omega}^2(|\mathbf{q}+\mathbf{G}|)} \tag{7.6}$$

$$\text{mit} \qquad \tilde{\omega}^2(q) \;=\; \omega_{\mathrm{Pl}}^2 \left(\frac{1}{\epsilon_\infty - 1} + \left(\frac{q}{q_{\mathrm{TF}}} \right)^2 \right) + \left(\frac{\hbar}{2m} \right)^2 q^4 \quad , \tag{7.7}$$

wobei q_{TF} die Thomas-Fermi-Wellenzahl ist. Für die Nichtdiagonal-Elemente verwenden wir einen von Falter et al. vorgeschlagenen Ausdruck:

$$\epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q},0) \;=\; \delta_{\mathbf{G},\mathbf{G}'} + \frac{\omega_{\mathrm{Pl}}^2}{A(1+b|\mathbf{q}+\mathbf{G}|^2)(1+b|\mathbf{q}+\mathbf{G}'|^2)} \times \frac{(\mathbf{q}+\mathbf{G})}{|\mathbf{q}+\mathbf{G}|} \frac{(\mathbf{q}+\mathbf{G}')}{|\mathbf{q}+\mathbf{G}'|} \frac{\rho(\mathbf{G}-\mathbf{G}')}{\rho(0)} \quad . \tag{7.8}$$

Um den Nenner $A(1+b|\mathbf{q}+\mathbf{G}|^2)(1+b|\mathbf{q}+\mathbf{G}'|^2)$ an den Nenner $\tilde{\omega}^2(q)$ des diagonalen Modells (7.6) anzupassen, muß der Parameter A wie folgt gewählt werden:

$$A = \omega_{\mathrm{Pl}}^2 \frac{1}{\epsilon_\infty - 1} \quad . \tag{7.9}$$

Dies entspricht dem gewünschten Verhalten $\epsilon(q,0) \rightarrow \epsilon_\infty$ für kleine q. Des weiteren wählen wir den Parameter b folgendermaßen:

$$b = \frac{\hbar}{2m} \frac{\sqrt{\epsilon_\infty - 1}}{\omega_{\mathrm{Pl}}} \quad . \tag{7.10}$$

Dadurch erhält der Nenner der von uns verwendeten Modellfunktion das gleiche Verhalten für große q (nämlich $\sim q^4$), wie es bei der diagonalen Modellfunktion (7.6) von Bechstedt et al. auftritt.

Im Si Volumensystem ergeben sich mit diesen beiden Modellfunktionen (B2) und (C2) Quasiteilchen-Bandstrukturen, deren Gapenergie ein wenig unter der der vollen RPA-Rechnung liegt. Ausgehend von der theoretischen Gitterkonstante (5.37 Å) und unter Benutzung der in RPA berechneten dielektrischen Konstante (ϵ_∞=13.0), ermitteln wir Gaps von 0.95 eV (Modell B2) sowie 1.09 eV (C2) im Vergleich zum RPA-Ergebnis von 1.23 eV.

Wir berechnen nun mit Hilfe dieser Modellfunktionen die Quasiteilchen-Energien der Dangling-Bond Zustände an der Si(001)-(2×1) Oberfläche. Die Berechnung erfolgt völlig analog zur Rechnung im letzten Abschnitt, in der wir die RPA-Funktion verwendet hatten (RPA-GWA). Wir stellen in Tabelle 7.1 die Ergebnisse vor und vergleichen sie mit den Daten der RPA-GWA.

Zur Beschreibung der dielektrischen Abschirmeigenschaften wählen wir für die dielektrische Konstante ϵ_∞ mit 12.3 einen etwas geringeren Wert als die makroskopische dielektrische Konstante des Volumenkristalls (vgl. Anhang E). Wir diskutieren

Bandstruktur-Energien für Si(001)-(2×1)

Diel. Funktion		LDA	GWA RPA	GWA B2	GWA B2	GWA C2	Exp.
	ϵ_∞			12.3	10.0	12.3	
$\mathbf{D_{up}}$	Γ	–0.30	–0.15	–0.25	–0.30	–0.15	–0.4[1],–0.1[2]
	J	–0.30	–0.20	–0.30	–0.35	–0.15	–0.4[1],–0.2[2]
	K	–0.85	–0.80	–1.00	–1.05	–0.85	–1.0[1]
	J'	–0.90	–0.85	–1.05	–1.05	–0.90	–1.0[1],–0.8[2]
	Bandbreite	0.95	1.00	1.05	1.05	1.05	0.6[1],0.8[2]
	mittl. Versch.		+0.15	–0.05	–0.10	+0.10	
$\mathbf{D_{down}}$	Γ	0.45	0.95	0.75	0.80	0.85	1.1[3]
	J	0.50	1.00	0.80	0.85	0.90	
	K	0.50	1.10	0.90	0.95	1.00	
	J'	0.20	0.85	0.65	0.70	0.70	
	Bandbreite	1.25	1.20	1.10	1.10	1.10	
	mittl. Versch.		+0.55	+0.35	+0.40	+0.45	
$E_{\rm gap}^{\rm ind}$	indir. Gap	0.20	0.70	0.65	0.70	0.55	0.44[3],0.64[4]

[1]Ref. [Uhr81] [2]Ref. [Joh90] [3]Ref. [Cha83] [4]Ref. [Mön81]

Tabelle 7.1: *Quasiteilchen-Energien der Dangling-Bond Zustände an der Si(001)-(2×1) Oberfläche (in eV). Erste Spalte: LDA. Zweite Spalte: GWA, basierend auf der in RPA voll berechneten dielektrischen Matrix. Dritte und vierte Spalte: GWA, basierend auf der dielektrischen Modellfunktion B2 von Hybertsen und Louie [Hyb88a], ausgewertet mit der dielektrischen Konstanten ϵ_∞=12.3 bzw. 10.0. Fünfte Spalte: GWA, basierend auf der dielektrischen Modellfunktion C2 von Bechstedt et al. [Bec81] sowie Falter et al. [Fal92, Fal93]. Sechste Spalte: Experimentelle Daten.*

zunächst die Bandstruktur-Ergebnisse, die wir auf der Basis der Modellfunktion (B2) erzielt haben. Anders als bei der RPA-GW Rechnung wird hier der D_{up} Zustand bezüglich der LDA ein wenig abgesenkt (besonders an K und J'), und zwar im Mittel um –0.05 eV. Die Anhebung von D_{down} beträgt +0.35 eV und ist somit geringer als der mit der RPA-Matrix berechnete Wert. Dieses Ergebnis war vor dem Hintergrund der geringeren Volumengap-Aufweitung durch das Modell (B2) auch zu erwarten. Dennoch ergibt sich mit 0.65 eV für das Oberflächen-Gap ein Wert, der

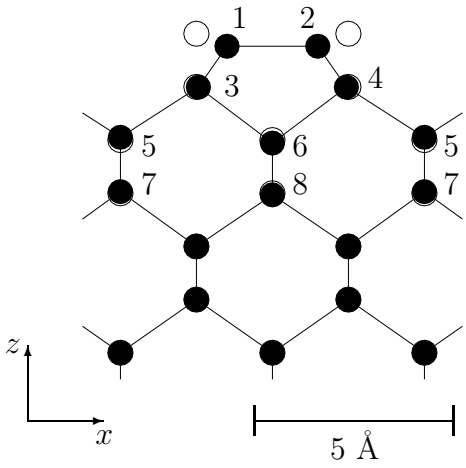
mit dem RPA-Resultat von 0.70 eV gut übereinstimmt. Das liegt an der niedrigeren Energie des D_{up} Zustands im Rahmen des Modells B2.

Die gleiche Modellfunktion (B2) wurde von Northrup zur Bestimmung der Quasiteilchen-Bandstruktur der Si(001)-$c(4\times2)$ Oberfläche benutzt [Nor93]. Dabei wurde die dielektrische Konstante zu 10.0 angenommen, um das Vakuum in der Superzelle zu berücksichtigen. Dieser Wert ist deutlich kleiner als die von uns verwendete dielektrische Konstante und als der experimentelle Wert (11.7). Um einen quantitativen Vergleich mit den Ergebnissen von Northrup durchführen zu können, haben wir unsere Rechnung mit dieser dielektrischen Konstante wiederholt. Als Folge der reduzierten Abschirmung erwartet man ein zusätzliches Absenken der besetzten und ein Anheben der unbesetzten Zustände. In unserer Rechnung wird der D_{up} Zustand um etwa –0.05 eV gegenüber der (ϵ_∞=12.3)-Rechnung abgesenkt und D_{down} um +0.05 eV angehoben. Das Gap wird um +0.05 eV vergrößert und beträgt nun 0.70 eV. [Im Volumenkristall beobachten wir die gleiche zusätzliche Gapaufweitung durch eine Reduktion der dielektrischen Konstante.] Im Rahmen dieses Modells (B2) mit ϵ_∞=10.0 lassen sich unsere Ergebnisse mit denen von Northrup vergleichen. Wir erzielen eine Quasiteilchen-Korrektur von –0.10 eV für die besetzten Dangling-Bond Zustände (Northrup: –0.15 eV) und von +0.40 eV für die unbesetzten (Northrup: +0.33 eV). Die Quasiteilchen-Korrektur für das Gap beträgt bei uns +0.50 eV und bei Northrup +0.48 eV.

Die Quasiteilchen-Korrekturen an den Oberflächen Si(001)-(2×1) und Si(001)-$c(4\times2)$ ähneln sich sehr. Das deutet darauf hin, daß die Eigenschaften der besetzten und der unbesetzten Dangling-Bond Zustände an den beiden Oberflächen nahezu gleich sind – es kommt lediglich an der $c(4\times2)$ Oberfläche zu etwas anderen Kopplungen und Aufspaltungen zwischen den Zuständen an benachbarten Dimeren.

Des weiteren haben wir eine GW-Rechnung mit der dielektrischen Modellfunktion (C2) durchgeführt. Hierzu wurde die dielektrische Konstante wieder zu 12.3 angenommen. Bezüglich der LDA ergibt sich eine Verschiebung des D_{up} Zustands um +0.10 eV zu höheren Energien. Der D_{down} Zustand wird um +0.45 eV angehoben. Das Gap zwischen den beiden Zuständen erfährt eine Korrektur von +0.35 eV und beträgt schließlich 0.55 eV; es fällt also etwas kleiner aus als in der RPA-Rechnung und in der (B2)-Modellrechnung. Kreß et al. haben ebenfalls Quasiteilchen-Korrekturen für die Oberflächenzustände an der Si(001)-(2×1) Oberfläche bestimmt [Kre94]. Die Autoren verwenden ein vereinfachtes GW-Schema und eine dielektrische Modellfunktion, die dem Modell (C2) sehr ähnelt. Die Autoren erhalten ei-

Geometrie der Si(001)-(2×1) Oberfläche: SDM



Atom	Δx [Å]	Δz [Å]
1	0.764	–0.316
2	–0.764	–0.316
3	0.050	–0.005
4	–0.050	–0.005
5	0.000	0.071
6	0.000	–0.076
7	0.000	0.041
8	0.000	–0.041

Abbildung 7.8: *Geometrie des symmetrischen Dimer-Modells (SDM) an der Si(001)-(2×1) Oberfläche [Krü95, KrüPr] (vgl. Abb. 7.2).*

ne Absenkung des D_{up} Zustandes um –0.03 eV am Γ-Punkt (bei uns: +0.15 eV), während D_{down} am Γ-Punkt um +0.80 eV angehoben wird (bei uns: +0.40 eV). Das fundamentale Gap beträgt 0.80 eV und ist damit wesentlich größer als das von uns berechnete (0.55 eV). Diese Abweichungen in den Ergebnissen sind eventuell auf systematische Unterschiede zwischen dem Zugang in [Kre94] und unserer Methode zurückzuführen.

Zusammenfassend stellen wir fest, daß die beiden dielektrischen Modellfunktionen (B2) und (C2) Quasiteilchen-Energien für Si(001)-(2×1) ergeben, die gut mit den Resultaten einer GW-Rechnung übereinstimmen, die auf der vollen RPA Matrix basiert. Die dielektrischen Modellfunktionen bieten nicht nur für Volumensysteme, sondern auch an Oberflächen die Möglichkeit, Quasiteilchen-Energien mit hoher Genauigkeit zu bestimmen. Dies eröffnet ein großes Anwendungspotential bei der Behandlung komplizierterer Rekonstruktionen mit größeren Einheitszellen.

7.4 Das symmetrische Dimer-Modell (SDM)

Zu Beginn dieses Kapitels hatten wir die Frage diskutiert, ob die Dimere an der Si(001) Oberfläche grundsätzlich gekippt sind (ADM) oder nicht (SDM). Vom Stand-

Si(001)-(2×1): Nicht-symmetrisierte Dangling Bonds des SDM

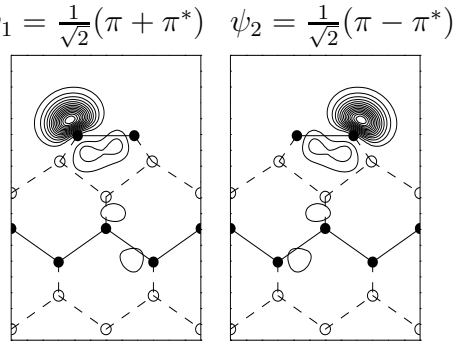


Abbildung 7.9: *Die an den einzelnen Dimeratomen lokalisierten Dangling Bonds $\psi_{1,2}$, aus denen die Zustände π und π^* des SDM linear zusammengesetzt sind. Die Darstellung erfolgt für den K-Punkt. [Die Berechnung von $\psi_{1,2}$ erfolgt ihrerseits aus den Zuständen π und π^* gemäß $\psi_1 = (\pi + \pi^*)/\sqrt{2}$ und $\psi_2 = (\pi - \pi^*)/\sqrt{2}$ unter Vernachlässigung des geringen Überlapps S.]*

punkt theoretischer Untersuchungen im Rahmen der LDA aus betrachtet, sprechen im wesentlichen zwei Gründe für das ADM. Zum einen ergibt sich durch das Kippen des Dimers ein Gewinn an totaler Energie. Zum anderen erweist sich die LDA-Bandstruktur des SDM als metallisch, was im klaren Widerspruch zum Experiment steht. Der Gewinn an totaler Energie durch das Kippen beträgt allerdings nur etwa 0.15 eV; vor diesem Hintergrund kommt dem zweiten Argument eine erhebliche Bedeutung zu. Es ist bekannt, daß die LDA das fundamentale Gap eines Halbleiters deutlich unterschätzt; daher ist der metallische Charakter innerhalb der LDA noch kein sehr starkes Argument gegen das SDM. Wir haben in den letzten Abschnitten gezeigt, daß die Energielücke zwischen den Dangling-Bond Bändern des ADM durch Quasiteilchen-Korrekturen um 0.5 eV vergrößert wird. Es erscheint daher durchaus möglich, daß auch die überlappenden Bänder des SDM durch Quasiteilchen-Korrekturen soweit getrennt werden, daß eine halbleitende Quasiteilchen-Bandstruktur entsteht. Um diese Vermutung zu untersuchen, berechnen wir im Folgenden die Quasiteilchen-Bandstruktur des SDM der Si(001)-(2×1) Oberfläche im Rahmen der GW-Näherung.

Wir stellen zunächst in Abb. 7.8 die von uns verwendete Oberflächen-Geometrie dar. Sie ergibt sich im Rahmen der LDA durch Minimieren der Gesamtenergie [Krü95] unter der Nebenbedingung, daß die Dimere symmetrisch sein sollen. Die

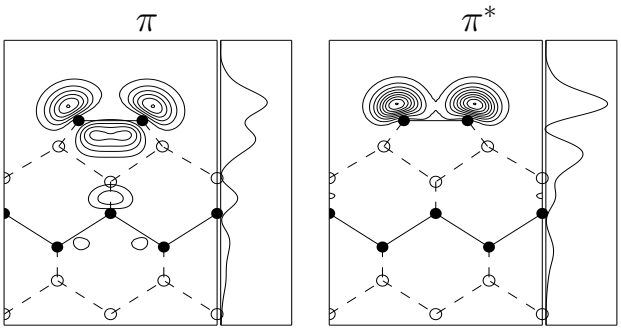


Abbildung 7.10: *Dangling-Bond Zustände im symmetrischen Dimer-Modell (SDM) der Si(001)-(2×1) Oberfläche (am K- Punkt). Der Zustand π weist generell eine tiefere Bandstruktur-Energie als der Zustand π* auf. Beide Zustände sind ähnlich stark an der Oberfläche lokalisiert.*

Dimere haben eine Bindungslänge von 2.27 Å.

An jedem Dimeratom bildet sich ein Dangling-Bond Zustand aus (siehe Abb. 7.9). Diese Zustände ähneln den beiden Dangling-Bond Zuständen D_{up} und D_{down} im ADM (vgl. Abb.7.3). Sie unterliegen nur einer sehr schwachen Wechselwirkung mit den übrigen Zuständen. In der elektronischen Struktur der Oberfläche sind diese beiden Dangling-Bond Zustände $\psi_{1,2}$ allerdings nicht direkt beobachtbar. Die symmetrische Geometrie des Dimers führt dazu, daß aus ψ_1 und ψ_2 symmetrische und antisymmetrische Linearkombinationen gebildet werden, die als π und π^* bezeichnet werden:

$$\pi = \frac{1}{\sqrt{2+2S}}(\psi_1 + \psi_2), \quad \pi^* = \frac{1}{\sqrt{2-2S}}(\psi_1 - \psi_2). \tag{7.11}$$

[Der Überlapp $S{=}\langle\psi_1|\psi_2\rangle$ zwischen ψ_1 und ψ_2 erweist sich mit 0.02 als sehr gering.] Diese beiden Zustände π und π^* ergeben sich als Eigenzustände in unseren Rechnungen. Ihre Ladungsdichten werden in Abb. 7.10 dargestellt. Hierin liegt der grundsätzliche Unterschied zwischen den elektronischen Strukturen des ADM und des SDM: Im ADM bleibt wegen der unterschiedlichen geometrischen Position des Up- und des Down-Atoms die räumliche Trennung der beiden Dangling-Bond Zustände erhalten: $\psi_1 \rightarrow D_{up}$, $\psi_2 \rightarrow D_{down}$. Damit ist auch eine energetische Separierung mit einem direkten Gap von 1-2 eV verbunden. Im SDM hingegen müssen sich aufgrund der Symmetrie des Dimers aus den beiden gleichwertigen Zuständen ψ_1 und ψ_2 die sym-

metrischen und antisymmetrischen Linearkombinationen π und π^* ausbilden. Die ursprünglich entartete Energie der Zustände $\psi_{1,2}$ spaltet bei diesem Vorgang auf:

$$E_{1,2} \rightarrow E^{\pm} = E_{1,2} \pm V \tag{7.12}$$

[unter Vernachlässigung des Überlapps S.] Die Bandstruktur der Dangling-Bond Zustände besteht also wie im Falle des ADM aus zwei Bändern. In unseren Rechnungen ergibt sich, daß diese beiden Bänder nahezu parallel verlaufen (siehe die Bandstruktur in Abb. 7.12). Die LDA-Bandbreite des unteren Bandes π beträgt 1.10 eV, die des π^* Bandes 1.25 eV. Die Aufspaltung $2V$ zwischen den Bändern ergibt sich zu 0.5...0.9 eV; die Kopplungskonstante V variiert also über die Brillouinzone nur sehr wenig. Anders als beim ADM sind die beiden Dangling-Bond Bänder nicht energetisch getrennt, sondern überlappen in der LDA-Bandstruktur um 0.65 eV. Die Wechselwirkung zwischen den Dangling-Bond Zuständen ist für eine energetische Separierung offenbar zu schwach.

Die energetische Auftrennung der Dangling-Bond Zustände in zwei Bänder hat beim SDM also eine völlig andere Ursache als im ADM. Daher ist die Vermutung, ihre Quasiteilchen-Korrekturen könnten für das SDM ähnliche Eigenschaften wie beim ADM aufweisen, a priori in keiner Weise gerechtfertigt. Wir wollen diese Quasiteilchen-Korrekturen im Folgenden untersuchen. Bei der dafür durchzuführenden GW-Rechnung sind einige Besonderheiten zu beachten, die sich aus dem metallischen Charakter der Bandstruktur ergeben. Bei dem halbleitenden ADM waren wir in der Lage, die dielektrische Matrix in RPA zu berechnen. Bei metallischen Systemen ist dies wesentlich aufwendiger. Der RPA-Ausdruck für die Polarisierbarkeit (3.12) enthält im Nenner Energiedifferenzen zwischen besetzten und unbesetzten Zuständen, die in der Nähe der Fermikante gegen Null gehen. Um den RPA-Ausdruck in dieser Situation konvergent zu bestimmen, ist eine äußerst aufwendige Brillouinzonen-Integration erforderlich. Wir bestimmen daher die dielektrische Matrix nicht im Rahmen der RPA. Statt dessen ziehen wir die dielektrische Modellfunktion (B2) von Hybertsen und Louie [Hyb88a] heran, die wir im letzten Abschnitt untersucht hatten. Auf das ADM angewendet, ergaben sich mit dieser Modellfunktion Quasiteilchen-Energien, die gut mit denen einer vollen RPA-Rechnung übereinstimmten. Die Abweichungen waren überall geringer als 0.2 eV. Wir erwarten daher, daß diese Modellfunktion auch für das SDM eine zuverlässige Quasiteilchen-Bandstruktur liefern wird. Die Modellfunktion erfordert die Eingabe der dielektrischen Konstanten ϵ_∞. Um einen direkten Vergleich zum ADM durchführen zu

können, verwenden wir auch für das SDM den gleichen Wert ϵ_∞=12.3. Selbst wenn die zu verwendende dielektrische Konstante für das SDM von der des ADM abweichen sollte (was wir hier nicht beurteilen können), so sind dadurch keine wesentlichen Änderungen in den Quasiteilchen-Energien zu erwarten. Bei einer Änderung um $\Delta\epsilon_\infty=\pm 2$ würden sich die Energien nur um ca. $\pm$0.05 eV verschieben (siehe vorigen Abschnitt).

Ein wichtiger Punkt betrifft die Frage nach der Besetzung eines jeden einzelnen Zustandes $|n\mathbf{q}\rangle$. Es sei daran erinnert, daß besetzte und unbesetzte Zustände in unterschiedlicher Weise in die Konstruktion des Selbstenergie-Operators und somit in die Berechnung der Selbstenergie eines Zustandes $|m\mathbf{k}\rangle$ eingehen (vgl. Gl. (4.12)). Beispielsweise besteht die Austauschenergie aus einer Summe nur über besetzte Zustände $|n\mathbf{q}\rangle$. Bei einem metallischen System durchschneidet die Fermifläche die Brillouinzone und trennt besetzte von unbesetzten Zuständen (siehe Abb. 7.11). Um in solch einer Situation den Besetzungszustand über die Brillouinzone genau verfolgen zu können, ist eine aufwendigere Integration mit mehr $\mathbf{q}$-Punkten erforderlich als bei einem halbleitenden System. Bei dem ADM hatten wir ein $\mathbf{q}$-Punkt-Schema mit zwei speziellen Punkten benutzt. Für das SDM verwenden wir acht spezielle Punkte. Ein Vergleich mit einer Rechnung mit zwei speziellen Punkten zeigt, daß mit acht speziellen Punkten alle Quasiteilchen-Energien bis auf etwa 0.1 eV konvergiert sind.

Die Frage nach der Besetzung ist noch in anderer Hinsicht wichtig. Ausgehend von der LDA-Bandstruktur (und der dadurch vorgegebenen Besetzung eines jeden Zustandes $|n\mathbf{q}\rangle$) berechnen wir im Rahmen der GWA neue Quasiteilchen-Energien. Wenn sich dadurch die Besetzung eines Bandes wesentlich ändert, wenn etwa ein im Rahmen der LDA besetzter Zustand nun unbesetzt ist, so hat man mit der einmaligen GW-Rechnung noch keine Selbstkonsistenz erreicht. Durch die Änderung der Besetzung tritt eine Veränderung im Selbstenergie-Operator ein, die wesentlich schwerer wiegt als etwa eine Änderung der eingehenden Bandstrukturenergien. In diesem Fall muß der gesamte Formalismus so lange iteriert werden, bis die Fermifläche selbstkonvergent feststeht. Wir stellen allerdings im vorliegenden Fall fest, daß die Besetzungsstruktur der Bänder sich durch die Quasiteilchen-Korrekturen *nicht* ändert. Die Dispersion der Bänder wird zwar geändert, und die GW-Bandstruktur weist eine höhere Fermi-Energie (0.62 eV) auf als die LDA-Bandstruktur (0.14 eV; vgl. Abb. 7.12). Auf die Einteilung der Bänder π und π^* in besetzte und unbesetzte Abschnitte hat dies aber nur einen marginalen Einfluß.

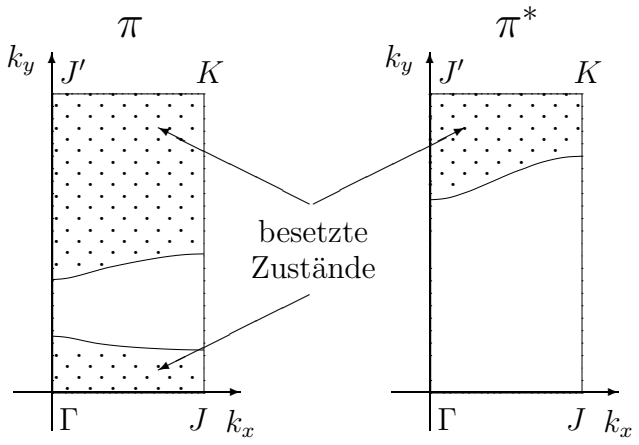


Abbildung 7.11: *Besetzung der Zustände π (links) und π* (rechts) im irreduziblen Teil der zweidimensionalen Oberflächen-Brillouinzone. Der schattierte Bereich gibt den jeweils besetzten Anteil der Brillouinzone an, definiert durch die Bedingung* $E_{\mathrm{m}\mathbf{k}} < E_F$. *Die Fermienergie beträgt 0.14 eV in LDA bzw. 0.62 eV in GWA. Die Besetzung der Zustände wird praktisch nicht verändert; sie ist in LDA und GWA gleich.*

Weitere GW-Iterationen sind daher nicht erforderlich. Wir stellen in Abb. 7.11 dar, in welchen Bereichen der Brillouinzone die Zustände π und π^* besetzt bzw. unbesetzt sind.

In Abb. 7.12 zeigen wir im unteren Teil die Quasiteilchen-Bandstruktur des SDM, zusammen mit der zugrunde liegende LDA-Bandstruktur (gestrichelt) und mit den experimentellen ARPES-Daten. Zum direkten Vergleich geben wir im oberen Teil der Abbildung nochmals die Bandstruktur für das ADM an (jetzt allerdings mit der dielektrischen Modellfunktion B2 berechnet; vgl. Spalte 3 in Tab.7.1). Das überwiegend besetzte π Band des SDM wird durch die Quasiteilchen-Korrekturen um +0.15...+0.50 eV zu höheren Energien verschoben. Seine Dispersion (1.10 eV in LDA) wird um 0.30 eV auf 1.40 eV erhöht. Das Band erfährt also eine völlig andere Quasiteilchen-Korrektur als das D_{up} Band des ADM, dessen Energie und Dispersion nahezu unverändert bleiben. Das überwiegend unbesetzte Band π^* wird um +0.30...+0.60 eV zu höheren Energien verschoben; seine Bandbreite vergrößert sich von 1.3 eV (LDA) um +0.2 eV auf 1.5 eV. Für das D_{down} Band des ADM hatten wir ähnliche Korrekturen festgestellt. Als wichtigstes Ergebnis unserer GW-Rechnung

für das SDM stellen wir fest, daß *die Quasiteilchen-Bandstruktur des SDM metallisch bleibt*. Die Quasiteilchen-Korrekturen verändern den metallischen Charakter der Bandstruktur nicht, und es besteht nach wie vor der eindeutige Widerspruch zu dem experimentellen Ergebnis, daß die Si(001)-(2×1) Oberfläche halbleitend ist. Die Ursache dafür liegt darin, daß die Quasiteilchen-Korrekturen zwar die Dispersion der Bänder π und π^* stark verändern, nicht aber ihre energetische Trennung $2V$. Diese hat sich gegenüber der LDA nur unwesentlich auf 0.6...1.0 eV erhöht; sie variiert nach wie vor nur sehr schwach über die Brillouinzone.

Die Quasiteilchen-Bandstruktur des SDM weicht deutlich von den experimentellen ARPES-Daten (einzelne Symbole in Abb. 7.12) ab. Die Dispersion des π Bandes ist zu groß, und das Band liegt um mehr als 0.5 eV über den experimentellen Werten. Die Quasiteilchen-Bandstruktur des ADM stimmt hingegen sehr gut mit den gemessenen Daten überein (siehe oberer Teil in Abb. 7.12 sowie die Diskussion in Abschnitt 7.2). Insbesondere erweist sich das ADM als halbleitendes System mit einer Energielücke von 0.65 eV (unter Verwendung der dielektrischen Modellfunktion B2), in guter Übereinstimmung mit der experimentell bestimmten Gap-Energie.

Im Folgenden diskutieren wir die Eigenschaften der Selbstenergie im SDM. In Abb. 7.13 stellen wir die Quasiteilchen-Korrekturen der Dangling-Bond Zustände (Kreuze) im Vergleich zu den entsprechenden Volumen-Daten (Punkte) dar. Die Abbildung entspricht Abb. 7.7 für das ADM.

Wir erkennen in Abb. 7.13 im wesentlichen den gleichen Trend wie bei den entsprechenden Daten für das ADM: Die Korrekturen der Dangling-Bond Zustände fallen generell schwächer aus als die der Volumen-Zustände. Die Korrekturdaten der Oberflächenzustände passen insgesamt sehr gut zu den Volumendaten. Die Stufe von 0.77 eV zwischen den Korrekturen besetzter und unbesetzter Volumenzustände wird deutlich reduziert, und die Gesamtheit aller Korrekturen ergibt eine glatte Funktion ohne Unstetigkeit. Es ist oftmals vermutet worden, die Unzulänglichkeiten der LDA ließen sich durch eine starre Verschiebung der Leitungsbänder zu höheren Energien, also durch die Anwendung eines „Scissors-Operators“, weitgehend überwinden. Das SDM bietet ein eindrucksvolles Beispiel für ein elektronisches System, bei dem dieses Verfahren nicht angewendet werden kann.

Ein wesentlicher Aspekt betrifft die Energie-Abhängigkeit der Korrekturen der Oberflächenzustände: Mit wachsender LDA-Bandstruktur-Energie steigt die Quasiteilchen-Korrektur an. Dieses Verhalten entspricht nicht der Energieabhängigkeit in einfachen metallischen Systemen. Beim homogenen Elektronengas stellt man fest,

Bandstrukturen für Si(001)-(2×1): ADM und SDM

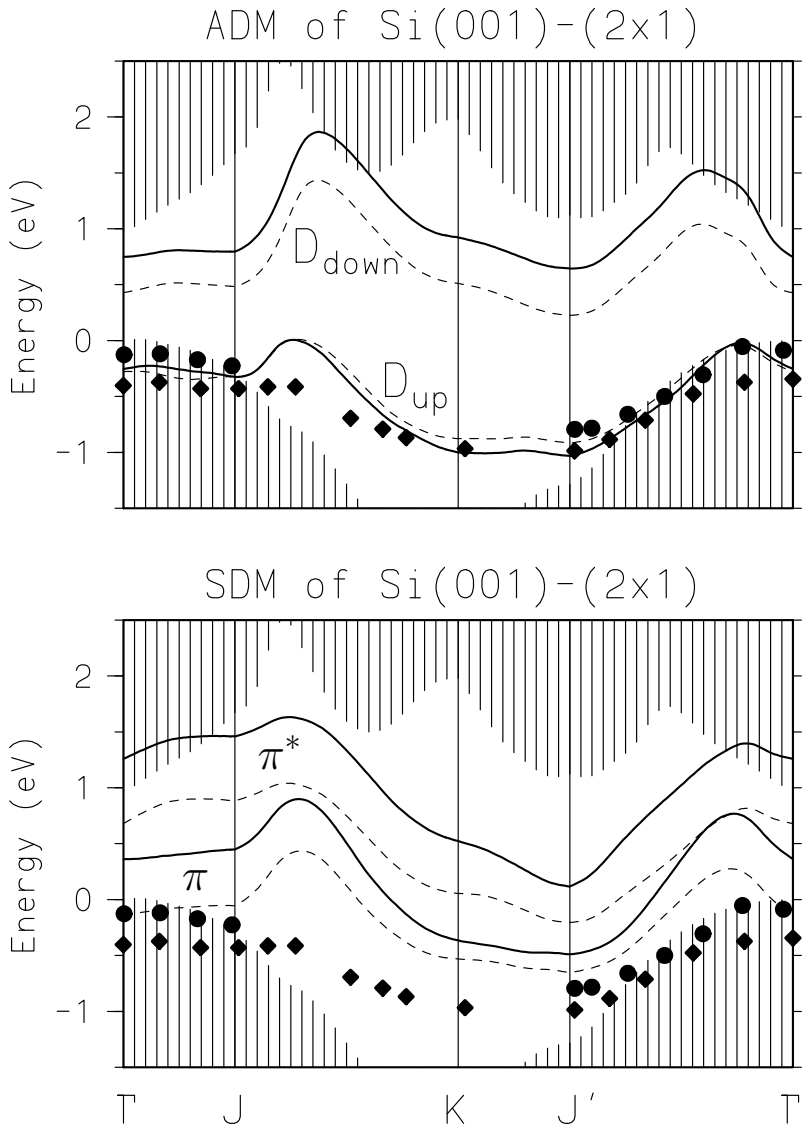


Abbildung 7.12: *Bandstruktur der Si(001)-(2×1) Oberfläche. Oben: für das ADM (vgl. Abschnitt 7.3). Unten: für das SDM. Wir vergleichen Quasiteilchen-Bandstrukturenergien (durchgezogene Kurven, berechnet in der GW-Näherung unter Verwendung der dielektrischen Modellfunktion B2) mit LDA-Energien (gestrichelt). Die einzelnen Symbole kennzeichnen experimentelle ARPES-Daten [Uhr81, Joh90].*

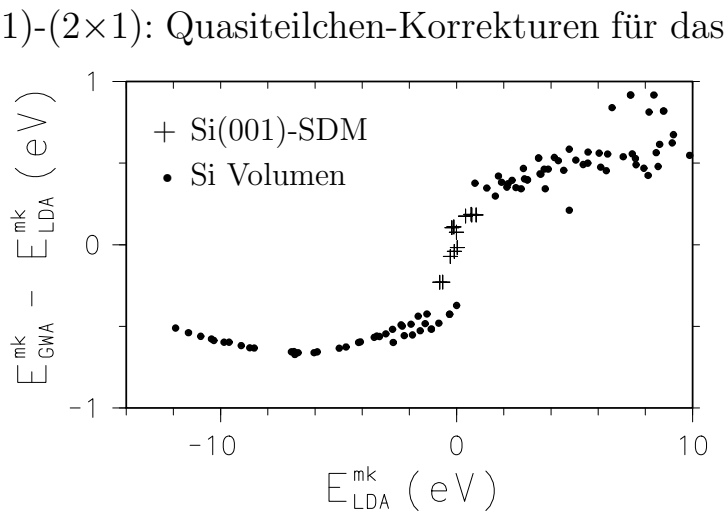


Abbildung 7.13: *Quasiteilchen-Korrekturen der Dangling-Bond Zustände π und π^* an der Si(001)-(2×1) Oberfläche im Rahmen des symmetrischen Dimer-Modells (SDM). Gleiche Darstellung wie in Abb. 7.7.*

daß die Quasiteilchen-Korrektur mit wachsender Energie *fällt* [Hed69]. Northrup et al. haben für die einfachen Metalle Li, Na und Al ebenfalls eine negative Energieabhängigkeit festgestellt [Nor89]. In dieser Hinsicht verhält sich das SDM nicht wie ein einfaches Metall oder wie ein homogenes Elektronensystem. Die Bänder π und π^* überlappen zwar, und die Bandstruktur ist metallisch. Das bedeutet aber nicht, daß an der Oberfläche freie Elektronen wie in einem einfachen dreidimensionalen Metall vorliegen. Abgesehen vom fehlenden elektronischen Gap in der Bandstruktur, entsprechen die Eigenschaften des SDM weitgehend denen der halbleitenden ADM-Oberfläche.

In Abb. 7.14 stellen wir die Beiträge zur Selbstenergie der SDM Dangling-Bond Zustände (Austausch-Anteil, Korrelations-Anteil) sowie die gesamte Selbstenergie dar. Die Notation ist die gleiche wie in Abb. 7.6 für das ADM. Hier werden nochmals die Zusammenhänge und Unterschiede zum ADM deutlich. Die Beiträge zur Selbstenergie der einzelnen Zustände folgen der durch die Volumen-Daten vorgegebenen Energie-Abhängigkeit. Anders als im ADM durchlaufen die SDM-Oberflächenzustände aber das gesamte Energieintervall des Gaps, und ihre Selbstenergie-Beiträge durchlaufen dementsprechend lückenlos sämtliche Werte zwischen denen der besetzten und unbesetzten Volumenzuständen.

Insgesamt zeigen das ADM und das SDM sehr ähnliche Selbstenergie- und Qua-

siteilchen-Korrektur-Eigenschaften. In beiden Systemen überdecken die Daten den gleichen Wertebereich. Die Werte des ADM sind an den Intervallgrenzen polarisiert, während die Werte des SDM das jeweilige Intervall gleichmäßig durchlaufen. Letztlich gehen die beiden Dangling-Bond Zustände D_{up} und D_{down} des ADM einerseits sowie π und π^* des SDM andererseits weitgehend als Linearkombinationen auseinander hervor. Daher erstaunt es nicht, daß die Erwartungswerte vieler Operatoren in diesen Zuständen sich wie entsprechende Linearkombinationen verhalten.

Neben der geringeren totalen Energie des ADM stellen wir als wesentlichen Unterschied zwischen dem ADM und dem SDM fest, daß die Bandstruktur des SDM metallisch ist, während das ADM ein halbleitendes System bildet, dessen Bandstruktur wesentlich besser mit allen experimentellen Befunden übereinstimmt. Aus diesem Unterschied leiten wir die Aussage ab, daß das symmetrische Dimer-Modell kein realistisches Bild der Si(001) Oberfläche bietet. Statt dessen besteht die Oberfläche aus asymmetrischen, gekippten Dimeren.

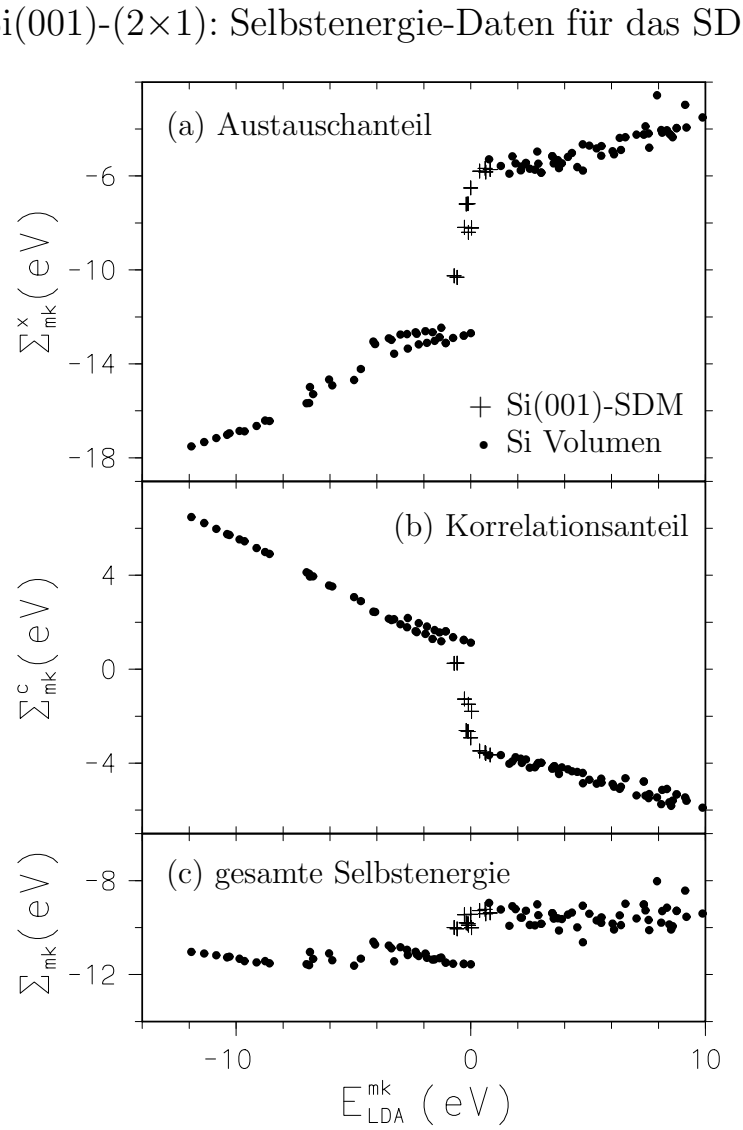


Abbildung 7.14: *Selbstenergien der Dangling-Bond Zustände π und π^* an der Si(001)-(2×1) Oberfläche im Rahmen des symmetrischen Dimer-Modells (SDM). Gleiche Darstellung wie in Abb. 7.6.*

Kapitel 8

Weitere reine Oberflächen

In diesem Kapitel untersuchen wir zwei weitere reine (001)-Oberflächen von Halbleitern der vierten Hauptgruppe, und zwar von Ge und von kubischem β-SiC (C-terminiert). Die Eigenschaften der beteiligten Elemente Ge bzw. C ähneln denen von Si. Die Atome besitzen jeweils vier Valenzelektronen und bevorzugen eine tetraedrische Konfiguration. Ge und β-SiC kristallisieren daher wie Si in der Diamantstruktur bzw. in der Zinkblendestruktur. An den idealen (001)-(1×1) Oberflächen dieser Kristalle befinden sich wie an der Si(001)-Oberfläche Atome, die nur noch an zwei Atome der zweiten Schicht chemisch gebunden sind. An jedem dieser Oberflächenatome liegen zwei Elektronen als unabgesättigte Bindungen (dangling bonds) vor. Die Umlagerung der Oberflächenatome innerhalb einer höheren Rekonstruktion als (1×1) ermöglicht die Bildung chemischer Bindungen zwischen ihnen. Damit ist ein Energiegewinn verbunden, und die Zahl der Dangling Bonds wird reduziert. Die zugehörigen elektronischen Dangling-Bond Zustände und Bänder stellen auch bei den Oberflächen Ge(001) und β-SiC(001) die auffälligsten Merkmale der elektronischen Struktur dar. Bei beiden Oberflächen wird durch diese Zustände die fundamentale Energielücke im Vergleich zum Volumengap reduziert.

8.1 Die Ge(001)-(2×1) Oberfläche

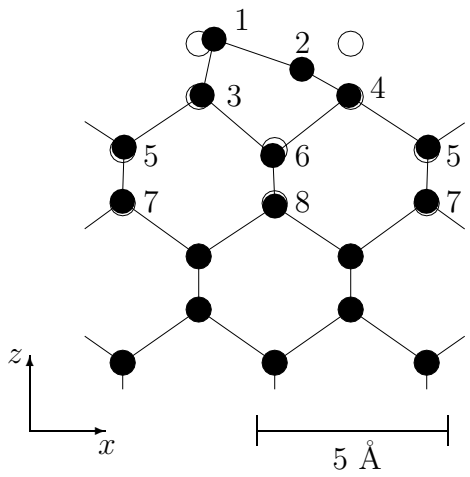
Die Dimerisierung der Oberflächenatome bildet das wesentliche Merkmal der Rekonstruktionsvorgänge an der reinen Ge(001) Oberfläche [Kub87]. Bei Raumtemperatur weist die Oberfläche eine nominale (2×1) Rekonstruktion auf [Kev85, Cul86, Kub87, Kip92, Lan94]. Die Intensitätsverteilungen von Ionenstreu-Experimenten [Cul86] und Röntgenbeugungsstudien [Ros92] deuten darauf hin, daß die Dimere

bezüglich der Oberflächenebene gekippt (asymmetrisch) sind, wie sie auch an der Si(001) Oberfläche beobachtet werden. Die Vorstellung von asymmetrische Dimeren wird durch das Auftreten höherer Rekonstruktionen bei Temperaturen unterhalb etwa 200 K, insbesondere der $c(4\times2)$ Rekonstruktion, unterstützt [Kev85, Cul86, Kub87, Lam87, Luc93]. Die Ursache für diese höhere Rekonstruktion liegt in der Absenkung der Gesamtenergie durch eine entgegengesetzte Kipp-Orientierung benachbarter Dimere [Nee88] (vgl. die Diskussion in Kap. 7). Die Intensitäten der entsprechenden Beugungsreflexe in LEED [Kev85], in der Helium-Streuung [Lam87] und in der Röntgenbeugung [Luc93] gehen oberhalb von 200 K mit steigender Temperatur zurück. Dieses Absinken läßt den Schluß zu, daß – genau wie bei Si(001) – ein Ordnungs-Unordnungs-Phasenübergang stattfindet [Nee88]. Bei Raumtemperatur ist nur noch eine nominale langreichweitige (2×1) Ordnung festzustellen. Die reine Ge(001) Oberfläche wurde im Rahmen von LDA-Rechnungen untersucht [Krü86, Pol87, Nee88, Lan90, Krü91, Krü94, Spi94, Krü95]. Alle Rechnungen finden übereinstimmend, daß die Oberfläche von asymmetrischen Dimeren bedeckt ist. Gekippte Dimere erweisen sich als energetisch günstiger als symmetrische Dimere. Der Energiegewinn durch das Kippen fällt mit 0.30 eV deutlich höher aus als an der Si(001) Oberfläche.

Wir konzentrieren uns hier auf die grundlegenden elektronischen Eigenschaften und die Quasiteilchen-Bandstruktur, die sich für solche gekippten Dimere ergeben. Dazu gehen wir von einer (2×1) Geometrie aus. Wir folgen der gleichen Vorgehensweise, mit der wir bereits das ADM der Si(001) Oberfläche behandelt haben. Der Slab besteht aus acht Schichten Ge. An der Oberseite befinden sich die Dimere. An der Unterseite belassen wir die Atome in ihren idealen Positionen und sättigen ihre freien Bindungen mit Wasserstoff ab. Die Beschreibung der Ge Atome erfolgt mit einem nichtseparablen Pseudopotential [Bac82]. Für die H-Atome verwenden wir wieder das von Gygi angegebene Pseudopotential [Gyg93]. Wir verwenden für Ge die Abklingkonstanten 0.19 und 0.50 im LDA-Basissatz. An den Oberflächenatomen vergrößern wir den Basissatz auf 30 Gauß-Orbitale (Abklingkonstanten: 0.16, 0.30 und 0.60), um eine höhere Flexibilität zu erreichen. Im GO-GWA Basissatz benutzen wir für die Ge Atome *s*-, *p*-, *d*- und s^*-Orbitale mit den Abklingkonstanten 0.19 und 0.50 und zusätzlich noch f-Orbitale mit der Abklingkonstante 0.50. Die Bestimmung der dielektrischen Matrix erfolgt in RPA (statische Matrix) mit einer Erweiterung zur dynamischen Matrix mit Hilfe des Plasmon-Pol Modells (I).

Als Oberflächengeometrie wird die in [Krü95] angegebene Geometrie zugrundege-

Ge(001)-(2×1): Oberflächengeometrie



Atom	Δx [Å]	Δz [Å]
1	0.399	0.119
2	–1.266	–0.671
3	0.089	0.046
4	–0.049	0.035
5	0.039	0.076
6	–0.049	–0.144
7	–0.008	0.050
8	0.006	–0.075

Abbildung 8.1: *Dimerbildung und Relaxation an der Ge(001)-(2×1) Oberfläche (vgl. Abb. 7.2) [Krü95]. Die offenen Kreise geben die idealen Positionen an.*

legt (siehe Abb. 8.1), die durch Energieminimierung im Rahmen der LDA bestimmt wurde. Der Kippwinkel nimmt mit 19° fast den gleichen Wert wie in Si an; dieser Wert stimmt mit experimentellen Befunden sehr gut überein [Cul86, Ros92]. Die Dimerbindungslänge ergibt sich in der LDA-Rechnung [Krü95] zu 2.42 Å und ist fast genau so groß wie die Bulkbindungslänge von 2.41 Å (basierend auf der theoretischen Gitterkonstante von 5.58 Å).

In der Elektronenstruktur der Oberfläche fallen besonders die beiden Dangling-Bond Zustände D_{up} und D_{down} auf. Sie werden hauptsächlich von den $4p_z$ Orbitalen der Ge Dimeratome gebildet. Für die meisten $\mathbf{k}_{\|}$-Vektoren aus der Oberflächen-Brillouinzone (SBZ) sind die Zustände deutlich an der Oberfläche lokalisiert (siehe linke und mittlere Figur in Abb. 8.2). Ihre räumlichen Ladungsdichte-Verteilungen gleichen denen der entsprechenden Dangling-Bond Zustände an Si(001) (vgl. Abb. 7.3). Auch an der Ge(001) Oberfläche kommt es im Zusammenhang mit dem Kippen des Dimers zu einer räumlichen Lokalisierung der beiden Zustände D_{up} und D_{down} an den Up-Atomen bzw. den Down-Atomen der Dimere. Die Bandstrukturenergien der Dangling-Bond Zustände liegen überwiegend außerhalb der projizierten Volumenbandstruktur (siehe Abb. 8.3), allerdings um etwa 0.4 eV tiefer als bei der Si(001) Oberfläche. Zwischen Γ und J sowie zwischen J' und Γ liegt der Zustand D_{up} im Bereich der besetzten Volumenzustände und befindet sich in Re-

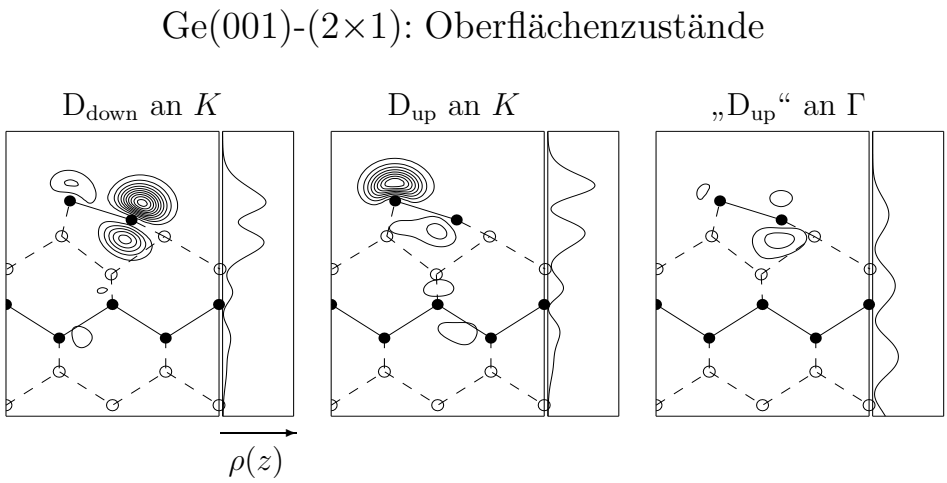


Abbildung 8.2: *Ladungsdichten der Dangling-Bond Zustände an der Ge(001)-(2×1) Oberfläche. Am* Γ *Punkt ist der oberste besetzte Zustand des Slabsystems (=Fortsetzung des* D_{up} *Bandes) über den ganzen Slab delokalisiert und nicht als Oberflächenzustand identifizierbar.*

sonanz mit ihnen. Seine Ladungsdichte ist nicht mehr an der Oberfläche lokalisiert, sondern verteilt sich nahezu gleichmäßig über den ganzen Slab (siehe rechte Figur in Abb. 8.2). Es ist nicht möglich, diesen Zustand im Bereich des Γ-Punktes als einen lokalisierten Oberflächenzustand zu interpretieren. Wir merken an, daß dieses Problem kein Artefakt der Slabmethode ist. Mit Hilfe streutheoretischer Methoden lassen sich Oberflächen von halbunendlichen Systemen behandeln [Pol80]. Solch ein Zugang ermöglicht auch die Untersuchung von räumlich weiter ausgedehnten Oberflächenresonanzen. Bei der Anwendung dieser Methode auf die Ge(001)-(2×1) Oberfläche stellt sich heraus, daß die lokale Zustandsdichte der Oberflächenatome am Γ-Punkt keinen lokalisierten Oberflächenzustand, sondern lediglich eine energetisch stark verbreiterte resonanzartige Struktur ohne ausgeprägte Lokalisierung aufweist [Krü86].

Bevor wir die Bandstruktur der Oberflächenzustände diskutieren, untersuchen wir zunächst die Abhängigkeit der Energien von der zugrundeliegenden Ge Volumen-Gitterkonstante. In der Volumen-Bandstruktur hängt die Größe der Energielücke sehr empfindlich von der Gitterkonstante ab. Bei der experimentellen Gitterkonstante von 5.65 Å erhalten wir für das direkte Gap am Γ-Punkt Ergebnisse von 0.01 eV in LDA und 0.59 eV in GWA. Die entsprechenden Werte bei der theoretischen Gitterkonstante von 5.58 Å betragen 0.37 eV in LDA und 0.98 eV in GWA. Das Gap

vergrößert sich also um 0.5 eV, wenn die Gitterkonstante um 0.1 Å reduziert wird. Die Vergrößerung des Gaps ist in LDA und GWA nahezu gleich groß. Das Volumengap tritt zwischen den 4*p*-artigen VBM-Zuständen und dem *s*-artigen untersten Leitungsband auf. Die zugehörige Energiedifferenz reagiert sehr empfindlich auf die Veränderung der Gitterkonstante. Die Energien der Dangling-Bond Oberflächenzustände relativ zum Volumen-VBM reagieren hingegen wesentlich schwächer auf die Variation der Gitterkonstante. Der Grund liegt darin, daß sowohl die Zustände des Volumen-VBM als auch die Dangling-Bond Zustände hauptsächlich von den Ge 4*p* Orbitalen gebildet werden. Der unterste Volumen-Leitungsbandzustand trägt zu den Dangling-Bond Zuständen praktisch nicht bei. Eine Veränderung der Gitterkonstante beeinflußt daher die Energien der Dangling-Bond Zustände relativ zu den Volumen-VBM Zuständen nur sehr wenig. Wir haben die LDA Oberflächen-Bandstruktur für beide Gitterkonstanten bestimmt und finden, daß die Energien der Dangling-Bond Zustände relativ zum jeweiligen Volumen-VBM sich in den beiden Rechnungen um weniger als 0.05 eV voneinander unterscheiden. Quantitative Aussagen über die Energien der Oberflächen-Zustände lassen sich somit bei beiden Gitterkonstanten gleichermaßen gut treffen. Wir führen im folgenden alle Rechnungen bei der theoretischen Gitterkonstante durch, bei der auch die optimale Oberflächengeometrie bestimmt wurde [Krü95].

In Abb. 8.3 vergleichen wir die Quasiteilchen-Bandstrukturenergien der Oberflächen-Zustände mit den zugrunde liegenden LDA-Energien (gestrichelt) und mit experimentellen Bandstruktur-Daten. Der Vollständigkeit halber schließen wir zusätzlich zu den beiden Dangling-Bond Zuständen D_{up} und D_{down} noch zwei weitere Oberflächenzustände in die Diskussion ein, deren Energien ebenfalls im Gap-Bereich der Volumenbandstruktur liegen. Der besetzte D_{up} Zustand erfährt nur eine sehr geringe effektive Quasiteilchen-Verschiebung von 0...+0.10 eV relativ zu den LDA-Werten. Seine Energie bleibt nahezu unverändert. Seine Dispersion (1.05 eV in LDA) wird um 0.10 eV auf 1.15 eV vergrößert. Diese Korrekturen entsprechen fast exakt den Effekten, die wir für den D_{up} Zustand an der Si(001)-(2×1) Oberfläche festgestellt hatten. Der besetzte Backbond-Zustand B erfährt ebenfalls nur eine sehr geringe Korrektur. Der unbesetzte D_{down} Zustand wird durch die Quasiteilchen-Korrekturen fast starr um +0.40...+0.50 eV zu höheren Energien verschoben. Diese Korrektur fällt etwas geringer aus als für den D_{down} Zustand an der Si(001)-(2×1) Oberfläche (+0.50...+0.65 eV). Wir merken an, daß bereits im Ge Volumensystem eine geringere Gapaufweitung auftritt (+0.61 eV) als in Si (+0.77 eV). Die Disper-

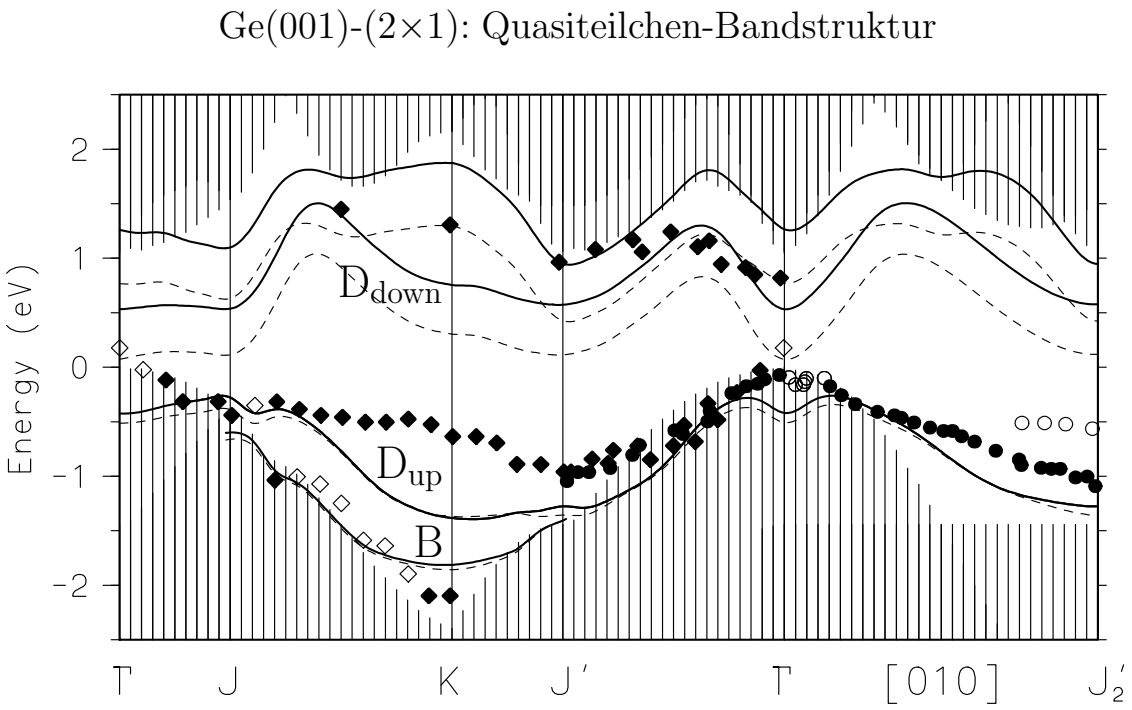


Abbildung 8.3: *Quasiteilchen-Bandstruktur der Oberflächen-Zustände an der Ge(001)-(2×1) Oberfläche. Wir vergleichen die GW-Ergebnisse mit den zugrunde liegenden LDA-Energien (gestrichelte Bänder) und mit experimentellen Daten (Punkte: ARPES-Daten von Landemark et al. [Lan90, Lan94], Rauten: ARPES- und ARIPES-Daten von Kipp et al. [Kip92, Kip95, Kip96] sowie von Skibowski und Kipp [Ski94]).*

sion des D_{down} Zustands beträgt sowohl in LDA als auch in GWA 1.00 eV. Die Korrektur des anderen unbesetzten Oberflächenzustands fällt ein wenig größer aus als die des D_{down} Zustands.

Wir vergleichen unsere berechneten Quasiteilchen Bandstruktur-Energien mit experimentellen ARPES-Daten von Landemark et al. [Lan90, Lan94] (Kreise) sowie mit ARPES- und ARIPES-Daten von Kipp et al. [Kip92, Kip95, Kip96] und Skibowski und Kipp [Ski94] (Rauten). Wir stellen in weiten Bereichen der SBZ eine gute Übereinstimmung zwischen den berechneten und den gemessenen Energien fest. Mit Ausnahme des D_{up} Zustands zwischen $JK/2$ und J' betragen alle Abweichungen weniger als 0.3 eV. Die gemessenen Daten des D_{up} Zustands liegen über den von uns bestimmten Energien. Die berechnete Dispersion des D_{up} Bandes stimmt entlang der $\Gamma J'$ Linie sehr gut mit der gemessenen Dispersion überein. Eine Ausnahme bildet die Umgebung des Γ Punktes. Hier stellen wir fest, daß das von uns berechnete Band

zum Γ-Punkt hin absinkt, während die experimentell gemessenen Energien zum Γ Punkt hin ansteigen. In den Messungen von Kipp et al. [Kip92] liegt die Energie des Bandes am Γ Punkt um 0.27 eV über dem VBM. Die Besetzung dieses Bandes wird in seinem Maximum am Γ-Punkt von der Fermienergie begrenzt; der Zustand verhält sich in diesem kleinen Bereich der SBZ metallisch. Dieser metallische Zustand wurde auch in anderen experimentellen Untersuchungen an der Ge(001) Oberfläche beobachtet [Kev84, Kev85, Zan89]. Aus unserer Rechnung erhalten wir keinen besetzten Zustand mit solch hoher Bandstrukturenergie. Wir stellen im Gegenteil fest, daß die Energie des besetzten D_{up} Zustands am Γ-Punkt um 0.45 eV unter dem Volumen-VBM liegt. Er koppelt resonant an die Volumenzustände an und ist nicht mehr an der Oberfläche lokalisiert (siehe oben).

Kevan und Stoffel [Kev84, Kev85] haben die Temperatur-Abhängigkeit der Photoemission dieses metallischen Oberflächenzustands untersucht. Sie stellten fest, daß Photoemission nur bei Temperaturen oberhalb 200 K auftritt. Bei tiefen Temperaturen ist der Zustand nicht nachweisbar. Die Autoren vermuten, daß das Auftreten des Zustandes durch temperaturbedingte dynamische Vorgänge in der atomaren Oberflächengeometrie hervorgerufen werden könnte. Wir erinnern daran, daß beispielsweise an der sehr ähnlich strukturierten Si(001) Oberfläche bei Raumtemperatur ein dynamisches Flippen der Dimere stattfindet [Dab92, Shk95]. Zu solchen Temperatur-Effekten erlaubt unser (T=0)-Formalismus keine Aussage.

Im Bereich zwischen $JK/2$ und J' stellen wir für den D_{up} Zustand deutliche Abweichungen zwischen unseren berechneten Energien und den experimentellen Daten von Kipp et al. [Kip92, Kip96] fest. Am K-Punkt beträgt die Differenz 0.8 eV. Bei der Energie –1.4 eV, die unsere Rechnung für den D_{up} Zustand am K-Punkt voraussagt, ist im Experiment keine Photoemission festzustellen [Kip96]. Rückfaltende Umklapp-Prozesse, wie wir sie bei der Interpretation der Si(001)-(2×1) Oberflächenbandstruktur in Abschnitt 7.2 diskutiert hatten, oder Zweidomäneneffekte vermögen diese Diskrepanz nicht vollständig zu erklären. Solche Effekte können zwar zusätzliche Strukturen im Photoemissionsspektrum am K-Punkt hervorrufen, aber sie können die eigentliche Photoemission des Zustands D_{up} nicht unterdrücken. Wir können an dieser Stelle lediglich mutmaßen, daß die im Experiment vorliegende Oberflächengeometrie möglicherweise anders ist als die, die sich aus der LDA Energieoptimierung ergibt. Es wäre allerdings erstaunlich, wenn ein Unterschied in der Geometrie einen solch starken Einfluß auf die Bandstruktur ausüben sollte. Selbst eine extreme Änderung wie etwa die Reduzierung des Kippwinkels auf Null führt in

der LDA Bandstruktur nur zu einer Anhebung des D_{up} Bandes am K-Punkt um 0.6 eV [Krü95]. Nahe des Punktes J_2' sind im Experiment Photoemissions-Strukturen bei etwa –0.5 eV zu erkennen [Lan94]. In unserer Rechnung stellen wir in diesem Bereich keine Zustände fest. Analog zur Si(001) Oberfläche könnte diese Emission auf Umklapp-Prozesse zurückzuführen sein, die durch lokale höhere Rekonstruktions-Korrelation hervorgerufen werden. Zusätzlich zum D_{up} Band beobachten Kipp et al. zwischen J und K bei –1...–2 eV einen weiteren besetzten Zustand [Kip92, Kip96], dessen Bandstrukturenergien sehr gut mit dem von uns berechneten Bandverlauf des Backbond-Zustands B übereinstimmen.

Im Bereich der unbesetzten Zustände stellen wir insgesamt eine zufriedenstellende Übereinstimmung zwischen den von uns berechneten Energien für den D_{down} Zustand mit den experimentellen ARIPES-Daten [Kip92, Ski94, Kip96] fest. Die gemessene Bandbreite von 0.6 eV ist jedoch geringer als unser Ergebnis von 1.00 eV. Zwischen Γ und J' wird experimentell eine geringere Dispersion festgestellt als in unserer Rechnung. Das Minimum der unbesetzten Zustände wird bei der Messung von Skibowski und Kipp [Ski94] am Γ-Punkt angenommen; es liegt um 0.85 eV über dem Volumen-VBM. Ortega et al. stellen hierfür in ihrer inversen Photoemissions-Untersuchung eine Energie von 0.6 eV über dem VBM fest [Ort93]. In unserer Rechnung befindet sich das Minimum des D_{down} Bandes ebenfalls am Γ-Punkt. Es hat eine Energie von 0.55 eV. Das ist zugleich die minimale Übergangsenergie zwischen besetzten und unbesetzten Zuständen. Die Oberflächen-Gapenergie wurde bisher nur in wenigen experimentellen Untersuchungen bestimmt. In Tunnelspektroskopie Experimenten wurde ein Oberflächengap von 0.9 eV festgestellt [Kub87]. Zum Tunnelstrom tragen hauptsächlich Oberflächenzustände in der Nähe des Γ-Punktes bei, da sie am langsamsten ins Vakuum hinein abklingen (vgl. Abschnitt 7.2) [Ter85]. Das Tunnelgap wird daher vom direkten Oberflächengap am Γ-Punkt bestimmt. Für diesen direkten Übergang zwischen den Dangling-Bond Zuständen am Γ-Punkt erhalten wir einen Wert von etwa 0.95 eV (dieser Wert ist wegen der erwähnten Lokalisierungs-Probleme am Γ-Punkt etwas unsicher), der gut mit dem Tunnelgap übereinstimmt. In der Messung von Kipp et al. konnten ARPES und ARIPES-Daten unter identischen Bedingungen an derselben Probe ermittelt werden [Kip96]. Daraus lassen sich direkt die Energiedifferenzen zwischen den besetzten und unbesetzten Oberflächenzuständen bestimmen. Die Autoren erhalten Energiedifferenzen zwischen 0.6 eV und 2.0 eV für direkte Übergänge. Diese Werte stimmen gut mit unseren Quasiteilchen-Energiedifferenzen von 0.95 eV bis 2.20 eV überein.

Ge(001)-(2×1) und Si(001)-(2×1): Quasiteilchen-Energien

		Ge(001)-(2×1)	Si(001)-(2×1)
D_{up}	mittl. QP Energie	–0.75	–0.30
	QP Bandbreite	1.15	1.00
	QP Korrektur	+0.05	+0.15
D_{down}	mittl. QP Energie	0.90	1.30
	QP Bandbreite	1.00	1.20
	QP Korrektur	+0.45	+0.55
Oberfl.-Gap	LDA Gap	0.40	0.20
	GWA QP Gap	0.80	0.70
	QP Korrektur	+0.40	+0.50
Volumen-Gap	QP Korrektur	+0.61	+0.77

Tabelle 8.1: *Berechnete Quasiteilchen (QP) Energiedaten der Dangling-Bond Zustände an den Oberflächen Ge(001)-(2×1) und Si(001)-(2×1) (in eV).*

8.2 Vergleich der Ge(001)-(2×1) mit der Si(001)-(2×1) Oberfläche

Ein Vergleich der Bandstruktur-Resultate mit den Ergebnissen aus Kap. 7 für die Si(001)-(2×1) Oberfläche läßt erkennen, daß die beiden Oberflächen sich in ihrer atomaren und elektronischen Struktur außerordentlich ähneln. Bei beiden Systemen bilden sich asymmetrische Dimere aus, die in Reihen angeordnet sind. Die Struktur der Dimere ist für Ge und Si nahezu identisch. Die Kippwinkel sind sehr ähnlich (19–20°). Die Dimerbindungslänge ist an der Ge(001) Oberfläche im wesentlichen durch die Ge Volumenbindungslänge gegeben. An der Si(001) Oberfläche ist sie um etwa 3 % kürzer als die Si Volumenbindung. Bei beiden Oberflächen kommt es bei Temperaturen unterhalb RT zu einem Phasenübergang von der (2×1) Phase zur $c(4\times2)$ Rekonstruktion. Aus dieser geometrischen Ähnlichkeit und aus der Analogie der elektronische Eigenschaften von Si und Ge resultiert eine große Ähnlichkeit der elektronischen Zustände an den beiden Oberflächen. In beiden Fällen treten im Bereich des fundamentalen Gaps zwei Dangling-Bond Zustände D_{up} und D_{down} auf, die an den Up- bzw. Down-Atomen der Dimere lokalisiert sind. In Tabelle 8.1

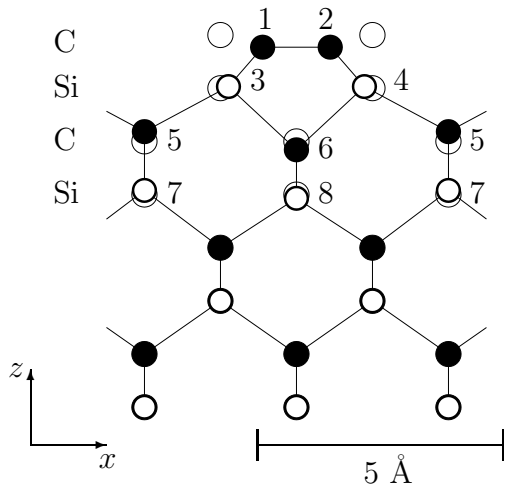
stellen wir einige charakteristische Energiedaten für die Dangling-Bond Bänder an den beiden Oberflächen zusammen. Die Bandbreiten der entsprechenden Zustände sind für Ge(001) und Si(001) bis auf 0.2 eV gleich groß, und sie weisen auch die gleiche Dispersion auf (man vergleiche Abb. 8.3 mit Abb. 7.4). Als wesentliche Unterschiede stellen wir fest, daß die Zustände an der Ge(001) Oberfläche um ca. 0.4 eV tiefer liegen als bei Si(001), und daß die Quasiteilchen-Korrekturen der unbesetzten Zustände und der Gapenergien bei Ge um 0.1 eV kleiner ausfallen als bei Si. Wir erinnern daran, daß bereits bei den Volumenkristallen die Gapkorrektur für Ge mit 0.61 eV niedriger ausfällt als für Si (+0.77 eV).

8.3 Die C-terminierte β-SiC(001)-(2×1) Oberfläche

Wir wenden uns nun dem Halbleitermaterial SiC zu. Während die Oberflächen der elementaren Gruppe-IV Halbleiter bereits sehr genau untersucht worden sind, beschäftigten sich bis vor kurzem nur relativ wenige Arbeiten mit den Oberflächen von SiC. Erst seit etwa fünf Jahren gewinnt SiC wegen seines enormen Anwendungspotentials an Bedeutung und wird sowohl experimentell als auch theoretisch intensiv erforscht. Für eine ausführliche Darstellung der äußerst komplexen Eigenschaften der verschiedenen SiC-Polytypen und ihrer Oberflächen verweisen wir auf die Arbeiten von M. Sabisch [Sab95, Sab96a, Sab96b] und die darin angegebenen Referenzen.

Wir wollen hier an einem Beispiel die grundlegenden elektronischen Eigenschaften von SiC-Oberflächen untersuchen. Wir wählen dazu die kubische Modifikation (β-SiC bzw. 3C-SiC), die unter den vielen Polytypen eine der strukturell einfachsten ist. Sie kristallisiert in Zinkblendestruktur. Die (001)-Oberflächen dieser Modifikation bestehen entweder aus einer reinen Si-Schicht oder aus einer reinen C-Schicht. Wir konzentrieren uns auf die letztere, also die C-terminierte Oberfläche. Die Struktur dieser Oberfläche wurde experimentell unter anderem von Powers et al. intensiv untersucht [Pow91]. Eine LDA-Studie niedrig indizierter Rekonstruktionen der Oberfläche favorisiert *in Reihen angeordnete symmetrische Dimere* als günstigste Geometrie für diese Oberfläche [Sab96a, Sab96b]. Diese Oberfläche weist eine (2×1) Rekonstruktion auf. In diesem Abschnitt soll nun die mit dieser Struktur verbundene elektronische Bandstruktur untersucht werden.

β-SiC(001)-(2×1) (C-term.): Oberflächengeometrie



Atom	Δx [Å]	Δz [Å]
1 (C)	0.852	–0.248
2 (C)	–0.852	–0.248
3 (Si)	0.162	0.033
4 (Si)	–0.162	0.033
5 (C)	0.000	0.199
6 (C)	0.000	–0.172
7 (Si)	0.000	0.084
8 (Si)	0.000	–0.077

Abbildung 8.4: *Die optimale Geometrie der C-terminierten β-SiC(001)-(2×1) Oberfläche besteht aus in Reihen angeordneten C=C Dimeren [Sab96a, Sab96b]. Die ausgefüllten Kreise stellen C-Atome, die offenen Kreise stellen Si-Atome dar. Die offenen Kreise im Hintergrund entsprechen den idealen Positionen.*

Wir stellen in Abb. 8.4 die optimierte Oberflächengeometrie vor. Der Dimer hat eine Bindungslänge von 1.36 Å. Zwischen den Dimeratomen bildet sich eine C=C Doppelbindung aus, wie sie etwa im C_2H_4 Molekül (Bindungslänge 1.34 Å) vorliegt. Der Bindungsabstand ist wesentlich kürzer als eine C–C Einfachbindung, die im C_2H_6 Molekül und in Diamant jeweils 1.55 Å lang ist. Die Geometrie des Dimers gleicht sehr dem Befund an der C(001) Oberfläche, an der sich ebenfalls symmetrische Dimere mit der gleichen Bindungslänge (1.37 Å) ausbilden [Krü95]. Offenbar wird die atomare Geometrie an der SiC Oberfläche von den C-Atomen dominiert; die Si-Atome in der zweiten Schicht der Oberfläche üben nur einen geringen Einfluß auf die Atompositionen in der Dimerschicht aus.

Der von uns verwendete Slab besteht aus vier SiC-Doppelschichten (insgesamt acht Atomlagen). An der Oberseite befinden sich die Dimere. An der Unterseite belassen wir die Atome in ihren idealen Positionen und sättigen ihre freien Bindungen mit Wasserstoff ab. In unseren Rechnungen beschreiben wir die Si- und H-Atome mit den gleichen separablen Pseudopotentialen wie in Kap. 7. Für die C-Atome verwenden wir ein von M. Sabisch [Sab93, Sab94] konstruiertes separables Pseudopotential. Der LDA-Basissatz besteht aus 20 Gauß-Orbitalen je Atom

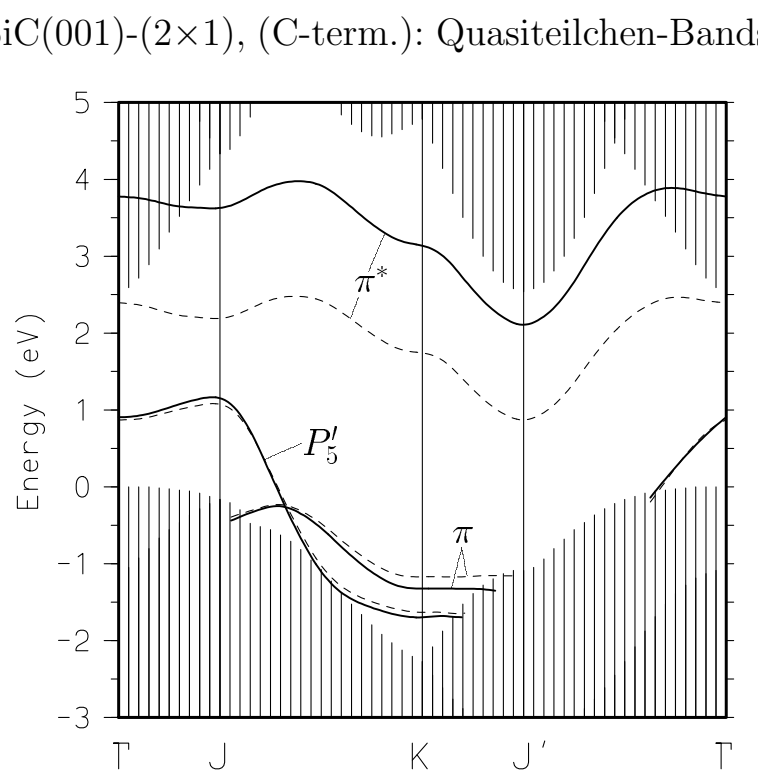


Abbildung 8.5: *Quasiteilchen-Bandstruktur der C-terminierten β-SiC(001)-(2×1) Oberfläche (in Reihen angeordnete C=C Dimere). Die gestrichelten Bänder stellen die LDA-Bandstruktur dar.*

(Abklingkonstanten: 0.2 und 0.6 für Si, 0.35 und 1.7 für C). An den C-Atomen in der Oberflächenschicht wird der Basissatz auf 30 Orbitale pro Atom erweitert (Abklingkonstanten: 0.25, 1.0 und 2.86). Der GWA-Basissatz besteht aus *s*-, *p*-, *d*- und s^*-Orbitalen (Abklingkonstanten: 0.2 und 0.6 für Si, 0.35 und 1.7 für C) sowie aus zusätzlichen *f*-Orbitalen (Abklingkonstante: 0.6 für Si, 1.7 für C). Die Bestimmung der dielektrischen Matrix erfolgt in RPA (Berechnung der statischen Matrix) mit einer Erweiterung zur dynamischen Matrix mit Hilfe des Plasmon-Pol Modells (I).

In Abb. 8.5 vergleichen wir die Quasiteilchen-Bandstruktur der Oberflächen-Zustände im Gapbereich mit der zugrunde liegenden LDA-Bandstruktur (gestrichelt). In dem betrachteten Energiebereich treten drei oberflächeninduzierte Zustände auf. Die beiden Dangling-Bond Zustände π und π^* (siehe Abb. 8.6) gehen als symmetrische und antisymmetrische Linearkombinationen aus den nichtabgesättigten Bindungen der beiden Oberflächenatome hervor. Sie bestehen hauptsächlich aus den C $2p_z$ Orbitalen der Dimeratome. Die direkten Übergangs-Energien zwischen

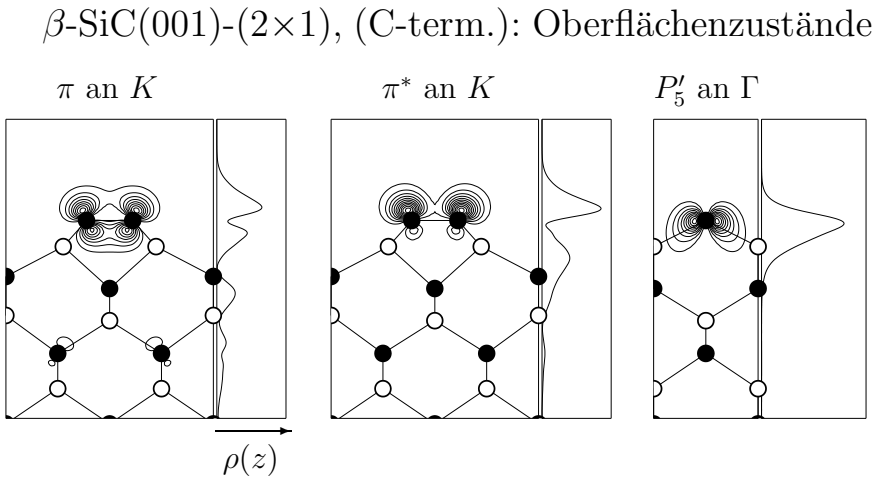


Abbildung 8.6: *Ladungsdichten der Oberflächenzustände an der β-SiC(001)-(2×1) Oberfläche, deren Energien im Gap-Bereich liegen. Die beiden Dangling-Bond Zustände π und π* werden in der Dimerebene dargestellt, der Zustand P_5' in der Ebene, die senkrecht dazu durch ein Dimeratom verläuft. Geschlossene Kreise markieren C-Atome, offene Kreise Si-Atome.*

π und π^* betragen 2.0 eV bis 3.0 eV in LDA. Sie sind wesentlich größer als die entsprechenden LDA-Energiedifferenzen zwischen den π und π^* Bändern im SDM an der Si(001)-(2×1) Oberfläche (0.5 bis 0.9 eV; vgl. Abschnitt 7.4); sie entsprechen vielmehr den direkten Übergangsenergien der analogen Zustände π und π^* an der C(001)-(2×1) Oberfläche (2.5 eV bis 3 eV in LDA) [Krü95]. Auch der Verlauf der Bänder ähnelt der Dispersion der π und π^* Bänder an der C(001)-(2×1) Oberfläche [Krü95, Kre94, Kre95]. Offenbar werden die Eigenschaften der Dangling-Bond Zustände durch die C Atome der Dimere festgelegt. Die Si Atome der zweiten Schicht üben auf die beiden Zustände π und π^* nur einen geringen Einfluß aus. Im Vergleich zu Diamant besitzt β-SiC ein wesentlich geringeres Gap. Das führt dazu, daß das π Band an der SiC(001)-(2×1) Oberfläche entlang der ΓJ Linie und entlang $J'\Gamma$ in der projizierten Volumenbandstruktur verläuft und resonant an volumenartige Zustände ankoppelt. In diesen Bereichen der SBZ kann der Zustand nicht als Oberflächenzustand interpretiert werden. An der C(001)-(2×1) Oberfläche hingegen liegen die Dangling-Bond Zustände durchgehend in der Energielücke der projizierten Volumenbandstruktur [Krü95].

Zusätzlich zu den beiden Dangling-Bond Zuständen π und π^* tritt in der Bandstruktur der β-SiC(001)-(2×1) Oberfläche noch ein weiterer Zustand auf (siehe Abb. 8.5), den wir in Anlehnung an [Sab96a] als P_5' bezeichnen. Dieser Oberflächenzu-

stand, der sich aus oberflächenparallelen C $2p$ Orbitalen der C Dimeratome zusammensetzt, stellt eine Besonderheit der SiC-Oberfläche im Vergleich zu den Oberflächen Si(001)-(2×1) und C(001)-(2×1) dar. Das Potential ist an den Si Atomen der zweiten Schicht wesentlich schwächer als an den C Atomen der Oberflächenschicht. Diese Barriere bewirkt sowohl die deutliche Lokalisierung des Zustands P_5' in der Dimerebene als auch seine ungewöhnlich hohe Bandstrukturenergie: sie liegt am J-Punkt um 1.1 eV über dem Volumen-VBM (in LDA). Durch diesen Zustand wird die Energielücke zwischen π und π^* (1.15 eV in LDA) nicht nur stark reduziert, sondern sogar geschlossen: Die LDA-Bandstruktur ist mit einem energetischen Überlapp von 0.2 eV zwischen den Zuständen P_5' an J und π^* an J' leicht metallisch. In der Realität jedoch sind praktisch alle Systeme, die aus Si und C zusammengesetzt werden können, halbleitend. Vor diesem Hintergrund stellt die Metallizität der LDA-Bandstruktur für die vorliegende SiC-Oberfläche ein erstaunliches Ergebnis dar. Um so interessanter ist die Frage, ob die Quasiteilchen-Bandstruktur dieses Systems ebenfalls metallisch ist (wie es beim SDM der Si(001) Oberfläche der Fall war), oder ob durch Quasiteilchen-Korrekturen die Metallizität überwunden wird, so daß sich eine halbleitende Bandstruktur ergibt.

Wenden wir uns daher nun der Berechnung der Quasiteilchen-Bandstruktur dieses Systems zu. Wir hatten in Abschnitt 7.4 bei der Untersuchung des SDM der Si(001)-(2×1) Oberfläche betont, daß die Besetzung eines jeden Zustands für die Konstruktion des Selbstenergie-Operators eine entscheidende Rolle spielt. Die Behandlung eines metallischen Systems wird dadurch wesentlich erschwert. Im Falle des SDM der Si(001) Oberfläche stellte sich glücklicherweise heraus, daß die Besetzungsstruktur der einzelnen Zustände durch die Quasiteilchen-Korrekturen praktisch nicht geändert wurden. Daher ergab sich bereits nach einer einzigen GW-Iteration ein konvergentes Resultat für die Quasiteilchen-Bandstruktur. Im Fall der β-SiC(001)-(2×1) Oberfläche liegt ein anderer Extremfall vor. Der metallische Überlapp der Bänder beträgt nur 0.2 eV. Wäre das System halbleitend, so könnte man auf der Basis der Erfahrungen an anderen halbleitenden Oberflächen erwarten, daß das π^* Band durch die Quasiteilchen-Korrekturen um mindestens 0.5 eV angehoben wird. Diese zu erwartende Anhebung übersteigt den in der LDA vorhandenen metallischen Überlapp. Vor diesem Hintergrund *nehmen wir zunächst an, daß das System halbleitend sein soll*. Konkret bedeutet diese Hypothese, daß wir den Zustand P_5' vollständig besetzen und π^* vollständig unbesetzt lassen. Unter dieser Annahme konstruieren wir den Selbstenergie-Operator und bestimmen die sich daraus ergebenden Quasiteilchen-

	ΔE_{gap} [eV]	
	Oberfläche	Volumen
Ge(001)	+0.40	+0.61
Si(001)	+0.50	+0.77
β-SiC(001)	+1.10	+0.98
C		+1.62

Tabelle 8.2: *Quasiteilchen-Korrekturen der Oberflächengaps der verschiedenen (001)-(2×1) Oberflächen im Vergleich zur Gapkorrektur im jeweiligen Substrat.*

Korrekturen und die Quasiteilchen-Bandstruktur (siehe Abb. 8.5). Die Energien der Zustände π und P_5' bleiben gegenüber der LDA praktisch unverändert, während das Band π^* um +1.2...+1.4 eV angehoben wird. Die Quasiteilchen-Bandstruktur ist halbleitend; sie weist ein indirektes fundamentales Oberflächen-Gap von 0.9 eV auf. Durch dieses Ergebnis ist die anfängliche Besetzungs-Hypothese a posteriori gerechtfertigt, und wir haben mit nur einer GW-Iteration ein konvergentes Resultst für die Besetzungsstruktur erzielt. Wären wir von der durch die LDA vorgegebenen Besetzung aus gestartet, so hätten wir letztlich das gleiche Ergebnis erhalten – allerdings erst nach mehreren GW-Iterationen.

Experimentelle Ergebnisse für die Energielücke an SiC(001) Oberflächen oder für die Bandstruktur-Energien von Oberflächenzuständen sind uns nicht bekannt. Auf der Basis unserer Untersuchungen prognostizieren wir, daß an der reinen C-terminierten β-SiC(001)-(2×1) Oberfläche besetzte elektronische Oberflächenzustände auftreten, deren Energien um bis zu 1.2 eV über dem Volumen-VBM liegen. Der unterste unbesetzte Oberflächenzustand hat eine Energie von 2.1 eV über dem VBM. Als minimale indirekte Energielücke erwarten wir einen Wert von 0.9 eV.

Die Quasiteilchen-Effekte an dieser Oberfläche gleichen denen, die wir an den Si(001)-(2×1) und Ge(001)-(2×1) Oberflächen gefunden hatten. Die besetzten Zustände erfahren nur sehr geringe Korrekturen, während die unbesetzten Bänder nahezu starr zu höheren Energien verschoben werden. Dadurch wird das Oberflächen-Gap vergrößert. In Tabelle 8.2 vergleichen wir die jeweiligen Quasiteilchen-Korrekturen der Gap-Energien miteinander. Die Gapaufweitung fiel für Si(001)-(2×1) und Ge(001)-(2×1) etwas *geringer* aus als die jeweilige Gapvergrößerung im

Si bzw. Ge Volumenkristall. Bei β-SiC(001)-(2×1) stellen wir hingegen fest, daß die Oberflächen-Gapaufweitung mit +1.1 eV (=Differenz zwischen dem Quasiteilchengap und dem negativen LDA-Gap) *größer* ist als die Aufweitung im β-SiC Volumenkristall (+0.98 eV). Das deutet wieder darauf hin, daß die Eigenschaften der Dangling-Bond Zustände an dieser Oberfläche durch den Dimer festgelegt werden – also durch Kohlenstoff. Die Gapaufweitung in Diamant ist mit +1.62 eV wesentlich größer als in β-SiC. Die Anwesenheit von Si in der zweiten Schicht der C-terminierten β-SiC(001)-(2×1) Oberfläche ist vergleichsweise weniger bedeutsam.

Kapitel 9

Adsorbat-bedeckte Oberflächen

Die reinen Oberflächen der Gruppe IV Halbleiter verhalten sich Fremdatomen gegenüber sehr reaktiv. Die nichtabgesättigten Bindungen (dangling bonds) der Oberflächenatome gleichen freien Radikalen, an denen Fremdatome aus dem Vakuum leicht angelagert werden können. Bei dieser Anlagerung von Fremdatomen werden die freien Bindungen in stabile chemische Bindungen umgewandelt. Dadurch verliert die Oberfläche ihre Reaktivität. Nach der Absättigung aller Dangling Bonds ist sie passiviert und verhält sich chemisch inert.

Durch die Adsorption von Adatomen und die Umwandlung der Dangling Bonds in chemische Bindungen verändern sich die Eigenschaften der Oberfläche massiv, denn die Atome, die ursprünglich die Oberfläche definierten, befinden sich nun in einer neuen chemischen Umgebung. Die modifizierte Elektronenstruktur führt zu einer Veränderung der Atompositionen. In vielen Fällen besitzt die adsorbatbedeckte Oberfläche eine andere Rekonstruktion als die reine Oberfläche. Die Anlagerung bestimmter Elemente an den reinen Oberflächen der Gruppe IV Halbleiter kann zu einem Aufbrechen der Oberflächendimere und einer Rückbildung der Rekonstruktion führen.

Die elektronische Bandstruktur erfährt durch die Adsorption ebenfalls gravierende Änderungen. Als Folge der Absättigung der freien Bindungen an den Oberflächenatomen treten die zugehörigen freien Dangling-Bond Zustände in der elektronischen Struktur nicht mehr auf. Statt dessen kommt es zur Ausbildung neuartiger elektronischer Oberflächenzustände, die den neu gebildeten chemischen Bindungen zuzuordnen sind. Die Untersuchung dieser Zustände wird im Mittelpunkt der folgenden Abschnitte stehen.

Ein häufiges Adsorbat ist Wasserstoff. H-Atome sättigen je eine freie Bindung ab.

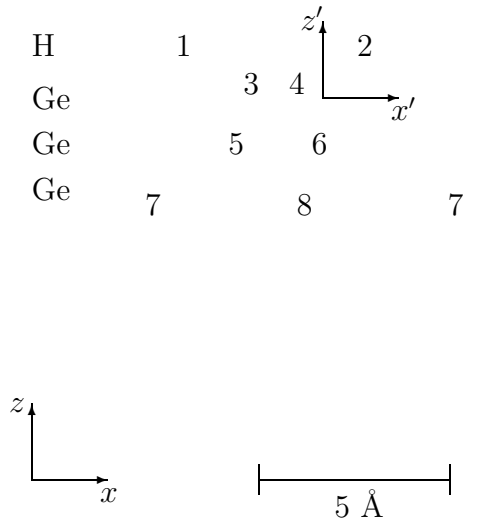
Sie besitzen eine sehr kleine räumliche Ausdehnung mit einem Kovalenzradius von nur 0.3 Å. Daher können auch an unrekonstruierten Oberflächen, die etwa im Falle der (001)-(1×1) Oberflächen zwei freie Bindungen pro Oberflächenatom aufweisen, alle freien Bindungen mit einzelnen H-Atomen terminiert werden. Als Beispiele für die Wasserstoff-Adsorption untersuchen wir die Oberflächen H:Ge(001)-(2×1) und H:Si(111)-(1×1).

Andere Adsorbate, wie etwa die Elemente der fünften und sechsten Hauptgruppe, bilden aufgrund ihrer elektronischen Konfiguration mit fünf bzw. sechs Valenzelektronen bevorzugt mehrere chemische Bindungen aus. Sie sättigen in der Regel mehrere freie Bindungen der Oberfläche gleichzeitig ab. Adatome der Gruppe VI können gezielt zur Terminierung von (001)-Oberflächen eingesetzt werden. Dabei wird im Falle einer vollständigen Bedeckung die ideale (1×1) Rekonstruktion wiederhergestellt. Den gleichen Zweck erfüllen Adatome der Gruppe V auf (111)-Oberflächen. Als Beispiel für solche Adsorbatsysteme behandeln wir die Oberflächen S:Ge(001)-(1×1) und As:Si(111)-(1×1).

9.1 Die H:Ge(001)-(2×1) Oberfläche

Es bestehen zwei Möglichkeiten, um die Ge(001) Oberfläche flächendeckend mit Wasserstoff abzusättigen.
(i) Die reine Ge(001) Oberfläche ist von asymmetrischen Ge Dimeren bedeckt (vgl. Abschnitt 8.1). Die Dangling Bonds an den Up-Atomen sind mit zwei Elektronen besetzt, die an den Down-Atomen sind unbesetzt. Bei der Anlagerung von Wasserstoff orientieren sich die Dimere parallel zur Oberfläche, und an dem halbbesetzten Dangling Bond eines jede Dimeratoms wird ein H-Atom angelagert. Das ermöglicht die Adsorption einer kompletten Monolage Wasserstoff. In dieser Konfiguration besteht die Oberfläche aus symmetrischen Ge Dimeren mit einem H-Adatom je Ge Dimeratom. Diese Monohydridphase weist eine (2×1) Rekonstruktion auf und wird als H:Ge(001)-(2×1) bezeichnet.
(ii) Wenn die Dimere an der Monohydrid-bedeckten Oberfläch aufgebrochen werden, so liegt an jedem Ge Oberflächenatom neben der bereits vorhandenen chemischen H-Ge Bindung noch eine weitere freie Bindung vor. An ihr kann ein weiteres H-Atom angelagert werden. Auf diese Weise kann eine weitere Monolage Wasserstoff adsorbiert werden. In dieser Dihydridphase liegt dann eine unrekonstruierte (1×1) Geometrie vor. Die Oberfläche wird als 2H:Ge(001)-(1×1) bezeichnet. Sie ist voll-

H:Ge(001)-(2×1): Oberflächengeometrie



Atom	x' [Å]	z' [Å]
1,2 (H)	±0.524	1.457
	Δx [Å]	Δz [Å]
3 (Ge)	0.740	–0.146
4 (Ge)	–0.740	–0.146
5 (Ge)	0.078	0.029
6 (Ge)	–0.078	0.029
7 (Ge)	0.000	0.072
8 (Ge)	0.000	–0.069

Abbildung 9.1: *Geometrie der Atome an der H:Ge(001)-(2×1) Oberfläche [KrüPr]. Der Ge-Ge Dimer ist symmetrisch. Die kleinen Punkte markieren H- Atome, die großen Kreise Ge-Atome.*

kommen passiviert und chemisch inert.

Experimentell läßt sich die Wasserstoff-Adsorption realisieren, indem die reine Ge(001) Oberfläche einer Atmosphäre aus atomarem Wasserstoff bei hoher Temperatur ausgesetzt und anschließend abgekühlt wird [Lan94]. Dabei bildet sich zunächst die Monohydridphase H:Ge(001)-(2×1). Diese Oberfläche ist noch nicht vollständig passiviert. Bei längerer Exposition der Oberfläche bei höherer Wasserstoff-Konzentration werden die Ge Dimerbindungen aufgebrochen, und es bildet sich die chemisch inerte Dihydridphase 2H:Ge(001)-(1×1) [Pap86].

Wir behandeln hier die Monohydrid-Oberfläche H:Ge(001)-(1×1), die von Landemark et al. mittels winkelaufgelöster Photoemission (ARPES) untersucht wurde [Lan94]. Die Minimierung der LDA-Gesamtenergie liefert für dieses System die folgende Oberflächengeometrie (siehe Abb. 9.1) [KrüPr]. An der Oberfläche befinden sich symmetrischer Dimere mit einer Dimer-Bindungslänge von 2.47 Å. An jedes Dimeratom ist je ein H-Adatom gebunden. Die H-Atome sättigen die freien Bindungen der Dimeratome ab und bilden chemische Ge-H Bindungen mit einer Bindungslänge von 1.55 Å. Der Winkel zwischen der Ge-H Bindung und der Ge-Ge Dimerbindung beträgt 109.8° und entspricht somit fast dem Tetraederwinkel von 109.5°. Die Ge-

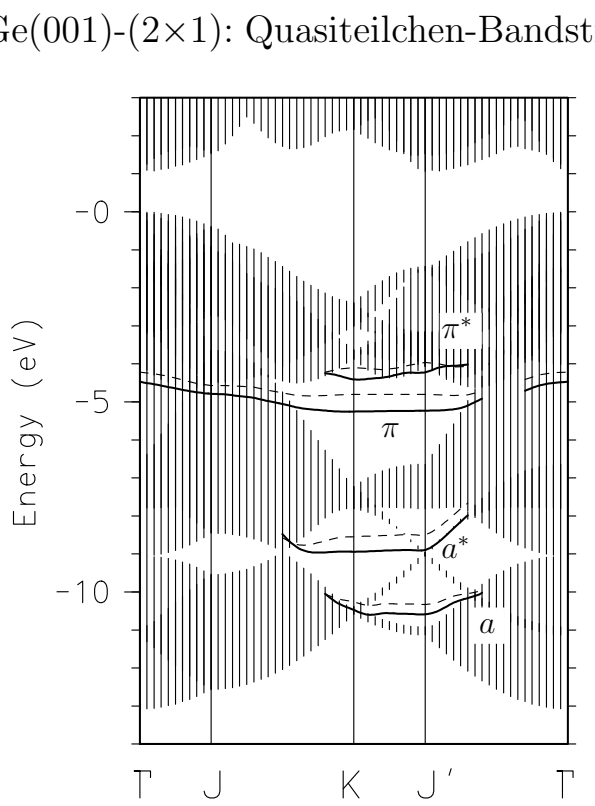


Abbildung 9.2: *Bandstruktur der H-induzierten Oberflächenzustände an der H:Ge(001)-(2×1) Oberfläche. Die durchgezogenen Linien markieren GW Quasiteilchen-Energien, die gestrichelten Linien bedeuten LDA-Energien.*

Dimeratome sind fast ideal tetraedrisch konfiguriert. Wir beschreiben die Oberfläche wieder mit einem Slabsystem. Der Slab besteht aus acht Schichten Ge. Wir konstruieren ihn so, daß er auf der Oberseite die zu untersuchende Oberflächengeometrie (Dimere mit H-Atomen) enthält. Die andere Oberfläche belassen wir wasserstofffrei, um eine Kopplung von H-induzierten Zuständen durch den Slab hindurch auszuschließen.

Die nichtabgesättigten Dangling Bonds werden durch die Wasserstoff-Absättigung in chemische Ge-H Bindungen überführt. Die Bandstruktur-Energien der zugehörigen Zustände liegen nicht mehr in der fundamentalen Energielücke, sondern werden durch das attraktive H-Potential zu tieferen Energien im Valenzband verschoben. In Abb. 9.2 stellen wir die elektronische Quasiteilchen-Bandstruktur vor, in der vier Oberflächenzustände auftreten. In Photoemissions-Messungen werden ebenfalls vier Oberflächen-Zustände beobachtet [App78]. Die beiden Zustände π und π^*

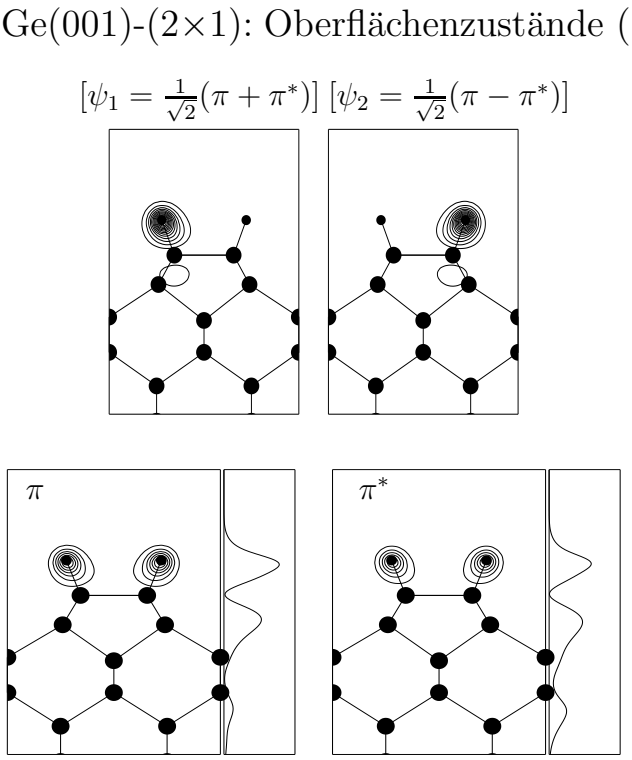


Abbildung 9.3: *Ladungsdichten der H-induzierten Zustände an der H:Ge(001)-(2×1) Oberfläche (am K-Punkt). Obere Zeile: Darstellung der Zustände $\psi_{1,2}$ der chemischen Ge-H Bindungen, die an den einzelnen H-Atomen lokalisiert sind, und aus denen die Zustände π und π^* linear zusammengesetzt sind. [Die Berechnung dieser Ladungsdichten erfolgte ihrerseits aus den Zuständen π und π^* unserer Slabrechnung gemäß $\psi_1 = 1/\sqrt{2}(\pi + \pi^*)$ und $\psi_2 = 1/\sqrt{2}(\pi - \pi^*)$.] Untere Zeile: Die symmetrische (π) und antisymmetrische (π^*) Linearkombination der Ge-H Bindungszustände.*

haben Quasiteilchen-Energien von –5.25 eV und –4.4 eV am J'-Punkt. Eine Analyse ihrer Wellenfunktionen zeigt, daß sie durch Hybridisierung der $1s$ Wellenfunktion der H-Adatome mit den $4p_z$ Valenzwellenfunktionen der benachbarten Ge Dimer-Atome entstehen. In Abb. 9.3 stellen wir in der unteren Zeile die Ladungsdichteverteilung der beiden Zustände π und π^* dar. Es handelt sich bei ihnen um symmetrische und antisymmetrische Linearkombinationen π und π^* von zwei Zuständen ψ_1 und ψ_2 der linken bzw. rechten Ge-H Bindung der Dimeratome (siehe obere Zeile in Abb. 9.3). Die Ausbildung gerader und ungerader Zustände wird durch die Symmetrie des Dimers erzwungen. Wir hatten den gleichen Effekt am SDM der Si(001)-(2×1) Oberfläche festgestellt und dort ausführlich diskutiert (vgl. Abschnitt 7.4). Die Aufspaltung $2V$ zwischen den beiden Bändern beträgt 0.70 eV (am K-Punkt). Der

H:Ge(001)-(2×1): Oberflächenzustände (2)

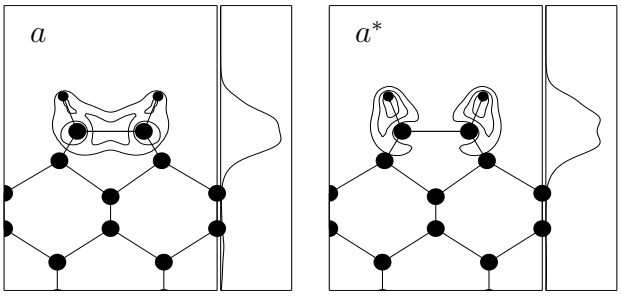


Abbildung 9.4: *Ladungsdichten der Zustände a und a^* an der H:Ge(001)-(2×1) Oberfläche (am Punkt $KJ'/2$). Der linke Zustand (als a bezeichnet) ist symmetrisch bezüglich der Spiegelebene des Dimers, der rechte (als a^* bezeichnet) ist antisymmetrisch.*

Überlapp S zwischen den beiden Zuständen ist nahezu Null. Das liegt an der ausgeprägten Lokalisierung der chemischen Ge-H Bindungen.

Landemark et al. [Lan94] haben die beiden Zustände π und π^* in ARPES Experimenten untersucht. Eine Analyse der Polarisationsabhängigkeit der Photoemission der beiden Zustände zeigt, daß der energetisch tiefere der beiden Zustände bezüglich des Dimers symmetrisch ist. Der andere Zustand ist antisymmetrisch. Wir finden in unseren Ergebnissen die gleiche Reihenfolge der Zustände.

Bei tieferen Quasiteilchen-Energien (–10.6 eV und –8.9 eV am K-Punkt) sind in der Bandstruktur zwei weitere Oberflächenzustände zu erkennen (siehe Abb. 9.4). Wir bezeichnen sie als a uns a^*. Auch diese beiden Zustände gehen als symmetrische und antisymmetrische Linearkombination aus zwei Zuständen hervor, die auf der linken bzw. auf der rechten Dimerseite lokalisiert sind. Sie stellen Kombinationen aus den $1s$ Wellenfunktionen der H-Adatome und der $4s$ Wellenfunktionen der benachbarten Ge Dimer-Atome dar. Da der Zustand $4s$ im Ge Atom eine niedrigere Energie als der $4p$ Zustand besitzt, erklärt sich hieraus die tiefere Energie der Zustände a und a^* im Vergleich zu π und π^*. Der Bereich des fundamentalen Gaps ist frei von Oberflächenzuständen. Die Oberfläche ist demnach halbleitend, und ihre Gap-Energie ist durch das Ge Volumengap gegeben.

Bevor wir die Quasiteilchen-Energien der Oberflächenzustände π und π^* mit experimentellen Ergebnissen vergleichen, untersuchen wir zunächst ihre räumlichen

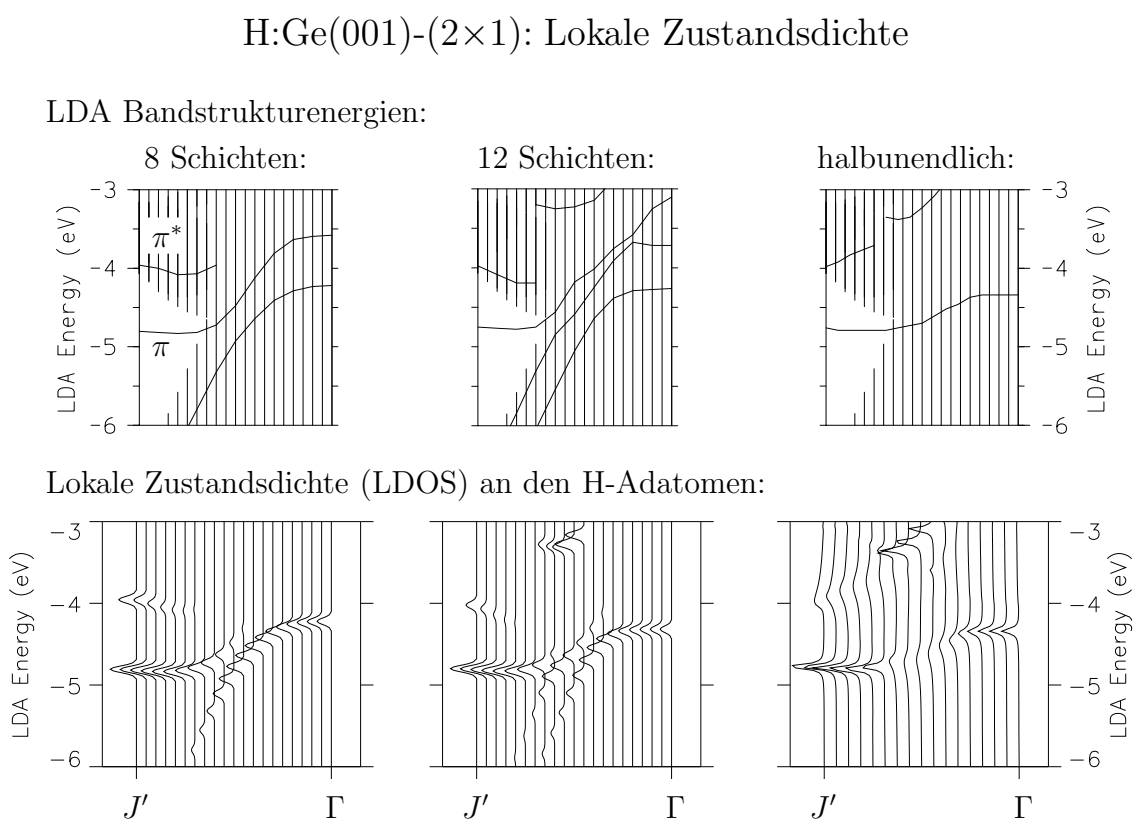


Abbildung 9.5: *LDA-Ergebnisse für die H:Ge(001)-(2×1) Oberfläche. Die Rechnungen wurden für zwei verschiedene Slabsysteme (8 bzw. 12 Schichten Ge) sowie mittels der Streutheoretischen Methode [Krü86] für ein halbunendliches System [KrüPr] durchgeführt. In der oberen Zeile werden die jeweiligen Bandstrukturenergien dargestellt, in der unteren Zeile die jeweils zugehörige* $\mathbf{k}_{\|}$*-aufgelöste lokale Zustandsdichte (LDOS) an den H-Adatomen.*

Lokalisierungs-Eigenschaften. Zu diesem Zweck stellen wir in Abb. 9.5 einen kleinen Ausschnitt aus der LDA-Bandstruktur entlang der $J'\Gamma$ Linie vor (obere Zeile). Zum Zwecke einer systematischen Studie wurde die Bandstruktur mit verschiedenen Slabdicken sowie für ein halbunendliches System [KrüPr] bestimmt. In der unteren Zeile stellen wir die entsprechenden Ausschnitte aus der $\mathbf{k}_{\|}$-aufgelösten lokalen Zustandsdichte (local density of states; LDOS) an den H-Adatomen vor. Wir beschränken uns jeweils auf die für die folgende Diskussion relevanten Bänder. Wir betrachten zunächst die bereits vorgestellte Rechnung mit dem 8-Schichten Slab (linke Seite in Abb. 9.5). Der π^* Zustand besitzt am J' Punkt ein starkes Gewicht an den H-Atomen, das jedoch in Richtung auf den Γ Punkt deutlich reduziert wird. Bereits auf der Mitte der $J'\Gamma$ Linie ist der Zustand in der LDOS nicht mehr zu erkennen. In

diesem Bereich koppelt er resonant an die Volumenzustände an und erstreckt sich räumlich über den gesamten Slab. In den ARPES Messungen von Landemark et al. nimmt die Photoemissions-Intensität dieses Zustands in analoger Weise ab, wenn der $\mathbf{k}_{\|}$-Vektor von J' in Richtung auf Γ verändert wird. Für den Zustand π stellen wir ein sehr interessantes Verhalten fest. Er läßt sich am J' Punkt als deutlich lokalisierter Oberflächenzustand erkennen. Die Dispersion des π Bandes ist in der Nähe des J' Punktes sehr schwach, und die Energie ist nahezu konstant. Mit dem Eintauchen in die projizierte Volumenbandstruktur zwischen J' und Γ erfährt das Band eine Dispersion zu höheren Energien, und sein Beitrag zur Zustandsdichte an den H-Atomen nimmt stark ab. Gleichzeitig erhält das nächsttiefere Band in Richtung auf Γ ein immer stärkeres Gewicht an den H-Atomen. Im Bereich des Γ-Punktes repräsentiert dieses Band den Zustand π. Im Bereich um $J'\Gamma/2$ hingegen ist der Zustand π kein exakter Eigenzustand des Slabsystems.

Zur Klärung der Frage, ob es sich hierbei um einen realen Effekt an der H:Ge(001)-(2×1) Oberfläche oder um ein Artefakt der Slabrechnung handelt, haben wir die gleiche Rechnung für einen dickeren Slab mit zwölf Ge-Schichten durchgeführt. In der Nähe des J'-Punktes und am Γ-Punkt verhält sich die Zustandsdichte des Zwölf-Schichten Slabs genau wie diejenige des Acht-Schichten Slabs. Der Zustand π besitzt einen ausgeprägten Anteil an den H-Atomen, und seine Energie zeigt fast keine Dispersion. Zwischen den beiden Bändern, die an J' und an Γ den π Zustand repräsentieren, tritt in dieser Rechnung ein weiteres Band auf, das in der Bandstruktur des Acht-Schichten Slabs nicht vorhanden war. In einem sehr engen Bereich der Brillouinzone nahe $J'\Gamma/2$ ist er an den H-Atomen lokalisiert und repräsentiert dort den Zustand π. Wir vermuten, daß bei weiterer Zunahme der Slabdicke immer mehr Bänder durch diesen Bereich der Bandstruktur hindurchlaufen werden, die jeweils für einen sehr engen $\mathbf{k}$-Raumbereich die Funktion des Zustands π übernehmen. Im Grenzfall sehr großer Schichtdicke erwarten wir im Spektrum des Systems einen kontinuierlichen Verlauf des Zustands π von J' bis Γ mit einer nur sehr schwach variierenden Energie. Diese Erwartung bestätigt sich im Vergleich mit einer analogen Rechnung für ein halbunendliches System im Rahmen der Streutheoretischen Methode [KrüPr]. In dieser Rechnung ergibt sich ein Spektrum, das tatsächlich einen kontinuierlich Verlauf des Zustands π aufweist (siehe Abb. 9.5). Im Bereich $J'\Gamma/2$ ist die Schärfe des Zustands allerdings stark reduziert, und seine Lokalisierung an den H-Atomen ist sehr schwach. Dies ist genau der Bereich in der Brillouinzone, in dem der von uns verfolgte Slabzugang Schwierigkeiten bei der Ausbildung des Zustands

zeigt. Die dortige schwache lokale Zustandsdichte des π Zustands an den H-Atomen liegt im Oberflächensystem begründet. Darüber hinaus tritt bei der Berechnung des Spektrums mit der Slabmethode das zusätzliche Artefakt in Erscheinung, daß der Zustand bei $J'\Gamma/2$ kein Eigenzustand des Systems ist. Das könnte daran liegen, daß an diesem Punkt der Zustand im realen System nur sehr langsam in das Substrat hinein abklingt. Solche Zustände werden durch die endliche Slabdicke und die resultierende Reflexion an der Slab-Unterseite zwangsläufig verfälscht. Der tatsächliche Verlauf der Bandstrukturenergie des Zustands π ist durch ein flaches Band mit sehr geringer Dispersion zwischen J' und Γ gekennzeichnet. Wir merken an, daß sowohl im 12-Schichten Slab als auch im halbunendlichen System im Bereich um $J'\Gamma/2$ ein weiterer Zustand bei höheren Energien (ca. –3 eV) an den H-Atomen auftritt. Er wird im Experiment [Lan94] nicht beobachtet.

Landemark et al. haben an der H:Ge(001)-(2×1) Oberfläche Bandstruktur-Untersuchungen mit winkelaufgelöster Photoemissions-Spektroskopie (ARPES) durchgeführt [Lan94]. Sie beobachten die beiden Oberflächenzustände π und π^*, die am J' Punkt die Energie –5.20 bzw –4.25 eV aufweisen. Zum Γ-Punkt hin steigt die Energie beider Zustände leicht an. Während der Zustand π bis zum Γ-Punkt verfolgt werden kann, nimmt die Photoemission aus dem Zustand π^* nach dem Eintritt in die projizierte Volumenbandstruktur so stark ab, daß er im Spektrum nicht mehr zu erkennen ist. Wir stellen in Abb. 9.6 die Quasiteilchen-Bandstruktur für die beiden Zustände π und π^* vor, zusammen mit den zugrunde liegenden LDA-Energien (gestrichelt) und mit den experimentellen Daten von Landemark et al. [Lan94]. Wir stellen eine hervorragende Übereinstimmung unserer Quasiteilchen-Energien mit den experimentellen Daten fest. Unsere Resultate von –5.25 eV und –4.20 eV am J' Punkt sind bis auf 0.05 eV mit den Meßergebnissen identisch. Die LDA-Energien liegen hingegen deutlich zu hoch. Der symmetrische π Zustand erfährt am J' Punkt eine Quasiteilchen-Korrektur von –0.45 eV, der antisymmetrische π^* Zustand von –0.30 eV zu niedrigeren Energien. Diese Quasiteilchen-Korrekturen weisen ein völlig anderes Verhalten auf als etwa für die Dangling-Bond Zustände an der reinen Ge(001)-(2×1) Oberfläche oder für Volumenzustände. Während bei allen bisher von uns untersuchten halbleitenden Oberflächensystemen die besetzten Oberflächen-Zustände praktisch keine Energieverschiebung erfahren, werden die H-induzierten Zustände um bis zu 0.45 eV abgesenkt, und zwar um so stärker, je ausgeprägter sie an den H-Adatomen lokalisiert sind. Das trifft besonders in solchen Bereichen der SBZ zu, in denen die Bänder außerhalb der projizierten Volumen-Bandstruktur liegen (etwa am

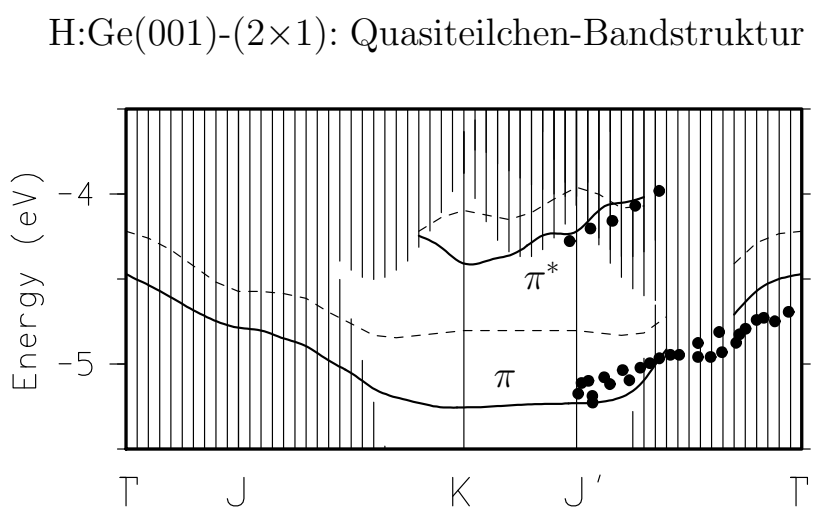


Abbildung 9.6: *GW Quasiteilchen-Bandstruktur der Zustände π und π^* an der H:Ge(001)-(2×1) Oberfläche. Die gestrichelten Bänder sind LDA-Energien. Die experimentellen ARPES-Daten wurden von Landemark et al. [Lan94] gemessen.*

J'-Punkt). Mit dem Eintritt in diese projizierte Volumen-Bandstruktur geht der π^* Zustand in einen volumenartigen Zustand über und verliert sein Gewicht an den H-Atomen (siehe oben). Dementsprechend wird seine negative Quasiteilchen-Korrektur reduziert. An denjenigen Punkten der SBZ, an denen er wie ein volumenartiger Zustand über den gesamten Slab delokalisiert ist, erfährt er auch nur eine schwache, volumenartige Korrektur. Dadurch wird seine Dispersion in diesem Übergangsbereich verändert. Das LDA-Band für π^* sinkt entlang $J'\Gamma$ ab, das Quasiteilchen-Band steigt hingegen an und stimmt hervorragend mit der experimentellen Dispersion überein. Der energetisch tiefere π Zustand ist über die gesamte SBZ gleichbleibend stark an den H-Adatomen lokalisiert (mit Ausnahme des Bereiches um $J'\Gamma/2$). Dementsprechend erfährt er auch überall fast die gleiche Absenkung um etwa –0.30 eV. Nur im Bereich um K und J' fällt sie ein wenig stärker aus (–0.45 eV). Die H-induzierten Zustände a und a^* erfahren ebenfalls deutliche Quasiteilchen-Korrekturen von –0.40 eV und –0.25 eV zu niedrigeren Bandstruktur-Energien.

Es ist schwierig, für die stark negative Quasiteilchen-Korrektur der H-induzierten Zustände eine einfache, plausible Begründung anzugeben. Sie unterscheiden sich von den anderen Zuständen des Systems hauptsächlich durch ihre sehr starke räumliche Lokalisierung auf einem sehr kleinen Raumbereich. Das freie H-Atom besitzt eine wesentlich kleinere Ausdehnung als andere Atome. Der 1s Zustand des H-Atoms er-

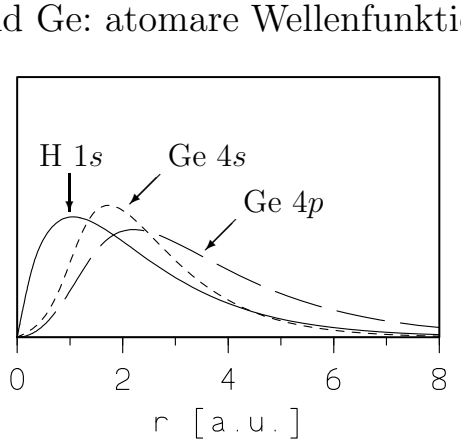


Abbildung 9.7: *Atomare Pseudo-Wellenfunktionen* $r\psi(\mathbf{r})$ *der Zustände 1s im H-Atom sowie 4s und 4p im Ge-Atom.*

reicht bereits bei a_0=0.53 Å das Maximum seiner radialen Ladungsdichteverteilung. Er ist also durch eine erheblich stärkere Konzentration und durch eine steilere Wellenfunktion gekennzeichnet als die Valenzzustände anderer Elemente. Wir stellen in Abb. 9.7 die 1s Wellenfunktion des H-Atoms den 4s und 4p Wellenfunktionen des Ge Atoms gegenüber. Daran wird die deutlich stärkere Lokalisierung des H 1s Orbitals sichtbar. Hier liegt ein möglicher Ansatzpunkt für die stärkere Quasiteilchen-Korrektur der H-induzierten Zustände an der H:Ge(001)-(2×1) Oberfläche im Vergleich zu den anderen Zuständen, die sich aus Ge Valenzorbitalen zusammensetzen. Das Austausch-Korrelations (XC) Potential der LDA reagiert wesentlich schwächer auf die Eigenschaften der einzelnen Wellenfunktionen, insbesondere auf ihre Lokalisierung. Es hängt im wesentlichen von der dritten Wurzel der Ladungsdichte ab,

$$V_{\mathrm{xc}} \sim \sqrt[3]{\rho} \sim (\psi)^{2/3} \quad , \tag{9.1}$$

während der Selbstenergie-Operator auch Abhängigkeiten in höherer Potenz zuläßt und als nichtlokaler Operator generell die Ortsabhängigkeit der Zustände besser erfaßt. Der Selbstenergie-Operator stellt einen wesentlich fundierteren Ausdruck für die Beschreibung der Austausch-Korrelations-Effekte dar als das LDA-XC Potential. Er liefert daher grundsätzlich einen genaueren Ausdruck für die Selbstenergien elektronischer Zustände als das LDA-Potential. Das gilt besonders, wenn sich die verschiedenen Zustände innerhalb eines Systems in ihren räumlichen Eigenschaften unterscheiden, wie etwa in unserem Falle der lokalisierte π^* Zustand am J' Punkt und

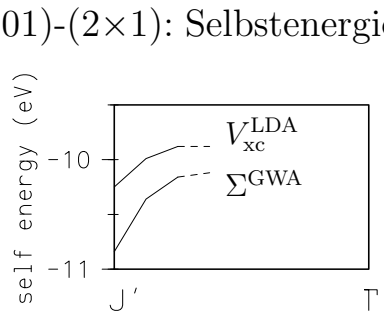


Abbildung 9.8: *Selbstenergie des Zustandes π^* an der H:Ge(001)-(2×1) Oberfläche. Die obere Kurve stellt den Erwartungswert des Austausch-Korrelations-Potentials der LDA, die untere Kurve den Erwartungswert des Selbstenergie-Operators der GWA dar.*

der volumenartig delokalisierte π^* Zustand im Bereich $J'\Gamma/2$. Um das zu verdeutlichen, stellen wir im interessanten Bereich zwischen J' und Γ die Erwartungswerte des LDA-XC Potentials und des Selbstenergie-Operators für den Zustand π^* dar (siehe Abb. 9.8). Das LDA-XC Potential ändert sich beim Übergang des Bandes vom lokalisierten H-Adatom Zustand zum volumenartigen Zustand nur um 0.37 eV, die GW-Selbstenergie jedoch um 0.72 eV. [Hierbei findet die stärkste Änderung im Austausch-Anteil statt.] Das führt in der Bandstruktur zu der deutlichen Änderung der Dispersion dieses Bandes.

Wir stellen in Abb. 9.6 eine sehr gute Übereinstimmung unserer Quasiteilchen-Energien mit den experimentellen Ergebnissen fest. Sowohl die Bandstrukturenergien der Bänder π und π^* als auch ihre Dispersionen werden durch unsere Rechnung sehr gut wiedergegeben. Offenbar liegt im Experiment genau die Oberflächengeometrie vor, wie wir sie auf der Basis der LDA Strukturuntersuchung für unsere Rechnung vorausgesetzt haben. Der GW-Formalismus ist in der Lage, auf dieser Geometrie aufbauend sehr zuverlässige Bandstrukturenergien für die stärker lokalisierten H-induzierten Zustände zu ermitteln.

9.2 Die H:Si(111)-(1×1) Oberfläche

Wir wenden uns nun mit der Si(111) Oberfläche einem System zu, das eine ganz andere Geometrie als die bisher diskutierten (001) Oberflächen aufweist. Während an den reinen (001)-Oberflächen der Elementhalbleiter jede mögliche Oberflächen-

Struktur freier (111)-(1×1) Oberflächen

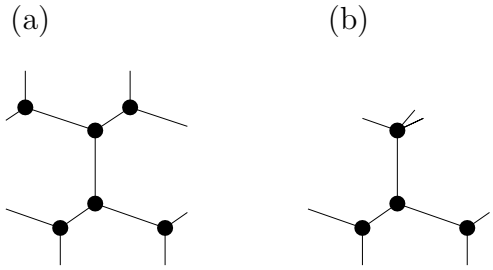
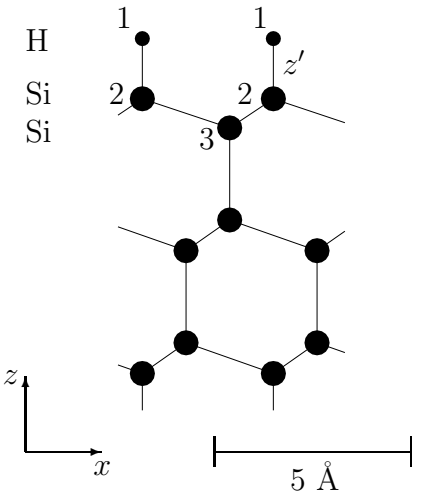


Abbildung 9.9: *Struktur freier idealer (111)-(1×1) Oberflächen von tetraedrisch gebundenen kubischen Halbleitern. Es sind zwei verschiedene Geometrien zu unterscheiden (a, b), je nachdem, an welcher Kristallebene der Kristall durchtrennt wird. In (a) weist jedes Oberflächenatom genau einen Dangling Bond auf, in (b) drei Dangling Bonds.*

schicht gleichwertig ist, müssen für die (111) Oberflächen zwei Fälle unterschieden werden (siehe Abb. 9.9). Im Fall (a) weist jedes Oberflächenatom eine freie Bindung auf (linkes Bild). Im Fall (b) besteht das System aus einer Monolage Substratmaterial weniger (oder mehr), und an jedem Oberflächenatom befinden sich drei freie Bindungen. Je nach Adsorbat wird die Anlagerung auf der (a) oder auf der (b) Oberfläche bevorzugt. Um die jeweils günstigste Oberflächengeometrie zu realisieren, kann das System durch reine Umlagerung in Form von Terrassenbildung von der Geometrie (a) in die Geometrie (b) bzw. umgekehrt umgewandelt werden.

Die Adsorption von Wasserstoff an Si(111) findet bevorzugt an der Oberfläche (a) statt. Die eine freie Bindung pro Si Oberflächenatom ermöglicht die Adsorption von einem H-Atom pro Oberflächenatom. Die Oberfläche weist dann eine (1×1) Rekonstruktion auf. Anders als an der Si(001) oder Ge(001) Oberfläche ist die Adsorption eines zweiten H-Atoms am Si-Oberflächenatom nicht möglich. Die Geometrie dieser H:Si(111)-(1×1) Oberfläche stellen wir in Abb. 9.10 vor. Durch die H-Absättigung weisen die Oberflächenatome eine fast ideale Tetraeder-Konfiguration auf. Die Atome befinden sich nahezu an ihren idealen Positionen; ihre Relaxationen sind äußerst gering [Bla94]. Wir berücksichtigen daher nur die Relaxation der äußersten Si-Atome. Eine Optimierung der LDA-Gesamtenergie ergibt hierfür eine Relaxation von 0.02 Å auf das Substrat zu. Parallel zur Oberfläche treten keine Relaxationen auf. In allen tieferen Schichten belassen wir die Si-Atome in ihren idealen Positionen.

H:Si(111)-(1×1): Oberflächengeometrie



Atom	z' [Å]
1 (H)	1.519
	Δz [Å]
2 (Si)	–0.019
3 (Si)	0.000

Abbildung 9.10: *Geometrie der H:Si(111)-(1×1) Oberfläche. Die kleinen Kreise markieren H-Adatome, die großen Kreise Si-Atome. Es treten nur Relaxationen in z-Richtung auf.*

Der Si-H Bindungsabstand ergibt sich zu 1.52 Å in guter Übereinstimmung mit dem Bindungsabstand im SiH_4 Molekül.

Durch die Absättigung der Dangling Bonds mit Wasserstoff bilden sich drei elektronische Oberflächen-Zustände aus, deren Energien wegen des attraktiven H-Potentials unterhalb des Gapbereichs liegen (siehe die Bandstruktur in Abb. 9.11). Sie sind besonders in solchen Bereichen der SBZ stark an den H-Atomen lokalisiert, in denen sie sich in Energie und Impuls von den volumenartigen Zustände unterscheiden, also außerhalb der projizierten Volumenbandstruktur liegen (vergleiche die Diskussion im vorigen Abschnitt). In Anlehnung an die Arbeit von Blase et al. bezeichnen wir sie als a, a' und b. Wir stellen in Abb. 9.12 die drei H-induzierten Zustände dar. Die Zustände a und a' kommen durch Hybridisierung der $1s$ Orbitale der H-Adatome mit den $3p_z$ Orbitalen der benachbarten Si-Atome zustande. Der Zustand b entspricht einer Überlagerung der H $1s$ Orbitale mit den benachbarten Si $3s$ Orbitalen.

Die Quasiteilchen-Korrekturen dieser drei Zustände gleichen denen, die wir an der H:Ge(001)-(2×1) Oberfläche für die Zustände π und π^* festgestellt hatten. Die LDA-Energien der Zustände liegen alle um 0.5 eV oder mehr über den experimentellen Daten, die von Hricovini et al. mit winkelaufgelöster Photoemission (AR-

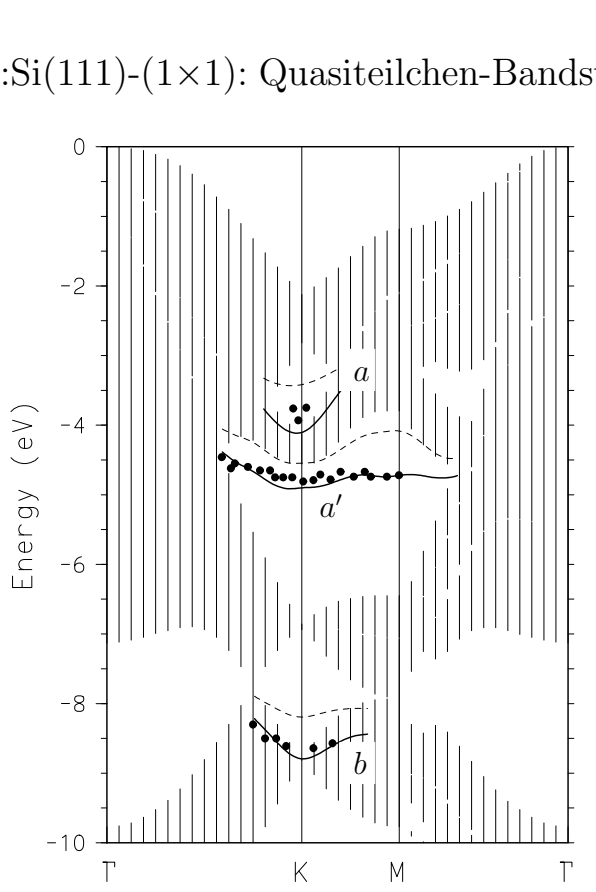


Abbildung 9.11: *Quasiteilchen-Bandstruktur der Wasserstoff-Zustände an der H:Si(111)-(1×1) Oberfläche im Vergleich zu experimentellen ARPES-Daten (Punkte) von Hricovini et al. [Hri93]. Die gestrichelten Linien stellen die LDA-Energien dar.*

PES) gemessen wurden [Hri93]. Alle drei Wasserstoff-Zustände erfahren durch die Quasiteilchen-Korrektur eine deutliche Energieabsenkung. Ihre GW-Quasiteilchen-Energien stimmen besonders für die Zustände a' und b hervorragend mit den experimentellen Daten überein. Wir stellen für den Zustand a' eine erhebliche Verbesserung der Dispersion des Bandes fest. Während die LDA-Bandstrukturenergien dieses Bandes entlang der KM Linie um 0.46 eV ansteigen, weisen die Quasiteilchen-Energien in diesem Bereich nur einen Anstieg um 0.17 eV auf, was mit den experimentellen Daten wesentlich besser übereinstimmt. Wir hatten im vorigen Abschnitt die analoge Verbesserung der Dispersion des π^* Bandes an der H:Ge(001)-(2×1) Oberfläche mit der unterschiedlich starken Oberflächenlokalisierung des Zustands für verschiedene $\mathbf{k}_\parallel$-Vektoren erklärt. Im Fall des a' Bandes an der H:Si(111)(1×1) Oberfläche liegt die gleiche Situation vor (siehe auch [Bla94]). Der a' Zustand ist am M-Punkt

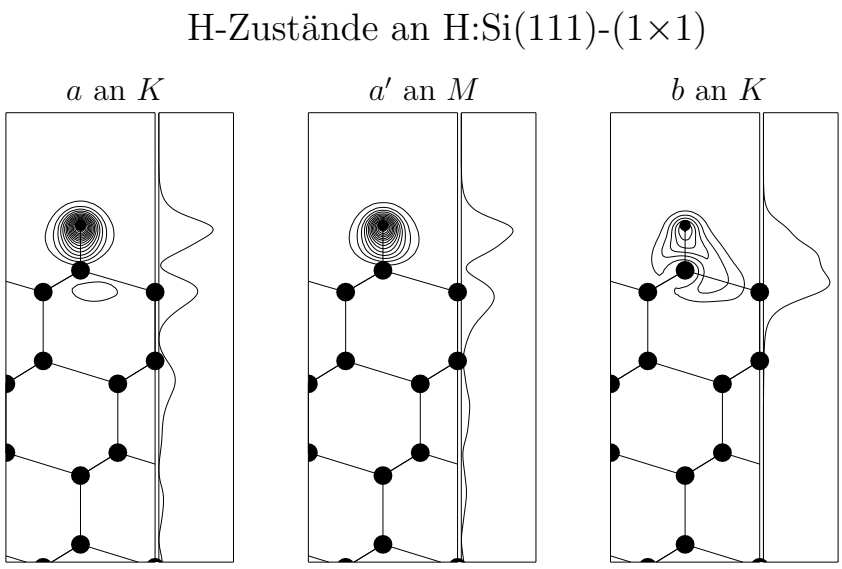


Abbildung 9.12: *Ladungsdichten der Oberflächenzustände an der H:Si(111)-(1×1) Oberfläche (am K-Punkt bzw. am M-Punkt).*

wesentlich stärker an den H-Adatomen lokalisiert als am K-Punkt, wo er an die volumenartigen Zustände ankoppeln kann und sich weiter über den Slab verteilt. Dementsprechend erfährt er am K-Punkt eine eher geringe Quasiteilchen-Korrektur von nur –0.35 eV, am M-Punkt hingegen eine stärkere Korrektur von –0.64 eV.

Blase et al. haben ebenfalls die GW-Quasiteilchen-Energien für die Wasserstoff-Zustände an der H:Si(111)-(1×1) Oberfläche berechnet [Bla94]. Unsere GO-GWA

H:Si(111)-(1×1): Quasiteilchen-Energien

[eV]	LDA	GWA	GWA [Bla94]	Exp. [Hri93]
a an K	–3.42	–4.10	–3.82	–3.80
a' an K	–4.55	–4.90	–4.76	–4.78
a' an M	–4.09	–4.73	–4.63	–4.76
b an K	–8.19	–8.79	–8.47	–8.64

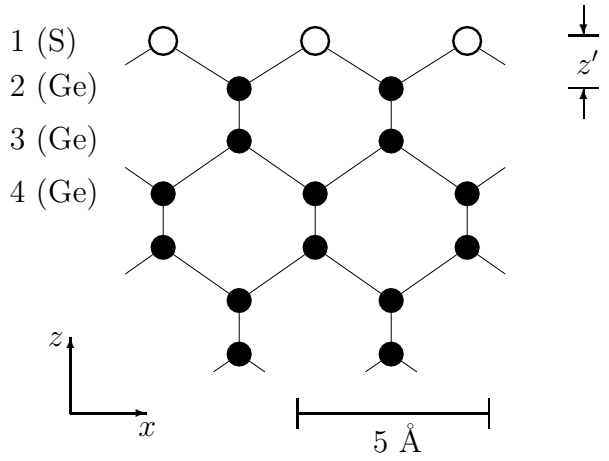
Tabelle 9.1: *Bandstruktur-Energien für die Oberflächenzustände an der H:Si(111)-(1×1) Oberfläche (in eV). Erste Spalte: LDA. Zweite Spalte: GO-GWA, basierend auf der in RPA berechneten statischen dielektrischen Matrix. Dritte Spalte: analoge GW-Rechnung, jedoch mit ebenen Wellen (PW-GWA) [Bla94]. Vierte Spalte: winkelaufgelöste Photoemissionsdaten (ARPES) [Hri93].*

Ergebnisse stimmen mit den PW-GWA Resultaten der Autoren gut überein (siehe Tabelle 9.1). Diese Übereinstimmung und die sehr gute Übereinstimmung der theoretischen mit den experimentellen Daten unterstreicht, daß unser methodischer Zugang zur Berechnung von Quasiteilchen-Energien im Rahmen der GO-GWA, basierend auf der in LDA optimierten Oberflächengeometrie, eine zuverlässige Beschreibung der H:Si(111)-(1×1) Oberfläche ermöglicht.

9.3 Die S:Ge(001)-(1×1) Oberfläche

Für die flächendeckende Adsorption von Schwefel auf der Ge(001) Oberfläche werden zwei mögliche Mechanismen beobachtet. Wenn wenig Schwefel zur Verfügung steht, so lagern sich S-Atome über den Ge-Dimeren der Oberfläche an [Leu88, New93]. Die Dimerbindungen werden aufgebrochen, und die S-Atome bilden chemische Bindungen mit den beiden Ge-Atomen, die ursprünglich den Dimer bildeten. Photoemission-Feinstruktur Untersuchungen [Leu88] deuten auf diese Brückenplatz-Adsorption hin. Auf diese Weise kann eine halbe Monolage Schwefel adsorbiert werden. Die Schwefelatome sind wie die ursprünglichen Dimere in Reihen angeordnet, und die Oberfläche weist eine (2×1) Rekonstruktion auf. Zwischen den Reihen bleiben weitere mögliche Adsorptionsplätze frei. Wenn der Oberfläche mehr Schwefel angeboten wird, so werden diese verbleibenden freien Adsorptionsplätze mit S-Atomen besetzt. Durch die Adsorption einer ganzen Monolage Schwefel kommt es zur Ausbildung der idealen (1×1) Geometrie. Die S Atome sättigen alle Dangling Bonds der Ge Oberflächenatome ab. Jedes Schwefelatom bildet dabei zwei chemische Bindungen mit Ge Atomen. Da Schwefel sechs Valenzelektronen besitzt, weist schließlich jedes S Atom eine Edelgas-Konfiguration aus acht Elektronen auf. Daher verhält sich die Adsorbat-bedeckte Oberfläche chemisch inert und ist vollständig passiviert. Der Adsorptionsprozeß läßt sich ohne Ausbildung von Störstellen realisieren, indem die Oberfläche mit elementarem Schwefel [Wes87] oder mit wäßriger $(NH_4)_2S$ Lösung [And95] in Kontakt gebracht wird.

Wir behandeln hier den Fall der vollständig passivierten S:Ge(001)-(1×1) Oberfläche. Für die mikroskopische Geometrie der Oberfläche kommen zwei mögliche Adsorptionsplätze für S in Betracht. Bei der oben erwähnten Anlagerung der S-Atome auf Brückenplätzen zwischen zwei Ge Oberflächenatomen bilden sich chemische S-Ge Einfachbindungen aus. Die Ge Oberflächenatome sind dann tetraedrisch koordiniert. Eine denkbare Alternative besteht in der Adsorption der S-Atome auf Plätzen direkt



Atom	z' [Å]
1 (S)	1.248
	Δz [Å]
2 (Ge)	–0.022
3 (Ge)	–0.001
4 (Ge)	0.010

Abbildung 9.13: *Geometrie der Atome an der S:Ge(001)-(1×1) Oberfläche [Krü90, Krü91, Pol92]. Die ausgefüllten Kreise markieren Ge-Atome, die offenen Kreise S-Atome.*

über den Ge Oberflächenatomen (Spitzenplatz-Position). Dabei müßten S=Ge Doppelbindungen gebildet werden. Bei der Untersuchung von teilbedeckten Oberflächen im Submonolagenbereich haben Weser et al. festgestellt, daß das Rumpfniveau-Spektrum der Ge $3d$ Elektronen mit der Brückenplatz-Geometrie verträglich ist, mit der Spitzenplatz-Geometrie jedoch nicht [Wes87]. Die Autoren schließen daher die Spitzenplatz-Geometrie als reale Adsorptionsgeometrie aus.

Die S:Ge(001)-(1×1) Oberfläche wurde theoretisch im Rahmen der LDA untersucht [Krü90, Krü91, Pol92]. Die Minimierung der totalen Energie favorisiert ebenfalls die Brückenplatz-Geometrie für die Adsorption der S Atome. Der Enegiegewinn gegenüber der Spitzenplatz-Geometrie beträgt 1.42 eV pro Einheitszelle. Wir stellen in Abb. 9.13 diese optimale Oberflächengeometrie vor. Die Relaxationen sind an dieser Oberfläche wesentlich geringer als an den Dimer-bedeckten reinen Ge(001) Oberflächen. Die Schwefel-Atome nehmen auf den Brückenplätzen fast die Positionen der Oberflächenatome an der idealen, nicht relaxierten Ge(001)-(1×1) Oberfläche ein. Die S-Adatome relaxieren ein wenig auf das Substrat zu, weil Schwefel zur dritten Periode des Periodensystems zählt und somit die Atome eine kleinere Ausdehnung besitzen als die Ge Atome aus der vierten Periode. Die Relaxation beträgt 0.15 Å. Die chemische S-Ge Bindung zu den Ge Oberflächenatomen ist dementsprechend mit 2.36 Å um etwa 2 % kürzer als die Ge-Ge Bindung des Ge Volumenkristalls.

Dieser Wert stimmt hervorragend mit den Ergebnissen von Photoemissions Feinstruktur Messungen von Leung et al. [Leu88] und mit SEXAFS Daten von Newstead et al. [New92] überein. Beide Untersuchungen finden eine S-Ge Bindungslänge von (2.36±0.05) Å an der S:Ge(001)-(2×1) Oberfläche, die mit einer halben Monolage Schwefel bedeckt ist.

Wir untersuchen im folgenden die elektronischen Eigenschaften der Oberfläche. Die dazu notwendigen LDA- und GW-Rechnungen führen wir genau so durch wie in den vorangegangenen Abschnitten. Für das erstmalig auftretende S-Atom benutzen wir in der LDA drei Abklingkonstanten (0.18, 0.57 und 0.9). Für den GW-Basissatz lauten die Abklingkonstanten 0.18, 0.5 und 0.8, mit zusätzlichen f-Orbitalen (0.5 und 0.8).

Die atomare Struktur der S:Ge(001)-(1×1) Oberfläche gleicht sehr der reinen, idealen, nicht rekonstruierten Ge(001)-(1×1) Oberfläche [Krü91]. Bezüglich der *elektronischen* Struktur ergeben sich hingegen wesentliche Unterschiede zur idealen Ge(001)-(1×1) Oberfläche, denn die adsorbierten S Atome weisen sechs Valenzelektronen auf anstelle der vier Valenzelektronen der Ge Atome. Wir stellen in Abb. 9.14 die Quasiteilchen-Bandstruktur der Oberfläche vor. Dabei konzentrieren wir uns auf den Energiebereich zwischen –6 eV und +3 eV. In diesem Bereich liegen vier lokalisierte Oberflächenzustände vor. Wir stellen in Abb. 9.15 die zugehörigen Ladungsdichten dar. Die Zustände D und D* bestehen hauptsächlich aus S $3p_z$ Orbitalen mit einem Beitrag der $4p_z$ Orbitale der benachbarten Ge Atome. Zusätzlich tritt ein Bridge-Bond Zustand Br auf, der von den S $3p$ Orbitalen in der Oberflächenebene gebildet wird. Ferner kommt es zur Ausbildung des Back-Bond Zustands B, der sich aus horizontalen S $3p$ Anteilen und aus $4p_z$ Orbitalen der benachbarten Ge Atome zusammensetzt. Die Zustände gleichen denen der reinen idealen Ge(001)-(1×1) Oberfläche [Krü91]. Sie weisen aber im vorliegenden Fall wegen des attraktiveren Potentials der S Adatome tiefere Bandstrukturenergien auf. Ferner besitzt S zwei Valenzelektronen mehr als Ge. Folglich werden die Zustände Br und D an der S:Ge(001)-(1×1) Oberfläche vollständig besetzt, während sie an der reinen Ge(001)-(1×1) Oberfläche nur teilweise besetzt sind [Krü91].

Der Vergleich der Quasiteilchen-Energien mit den LDA-Energien (gestrichelt) zeigt, daß die besetzten Bänder nahezu unverändert bleiben (siehe Abb. 9.14). Sie erfahren nur geringe Quasiteilchen-Korrekturen von maximal –0.2 eV zu niedrigeren Energien. Der unbesetzte D* Zustand wird nahezu starr um +0.45...+0.55 eV zu höheren Energien verschoben. Am Γ-Punkt stellen wir in der LDA-Bandstruktur

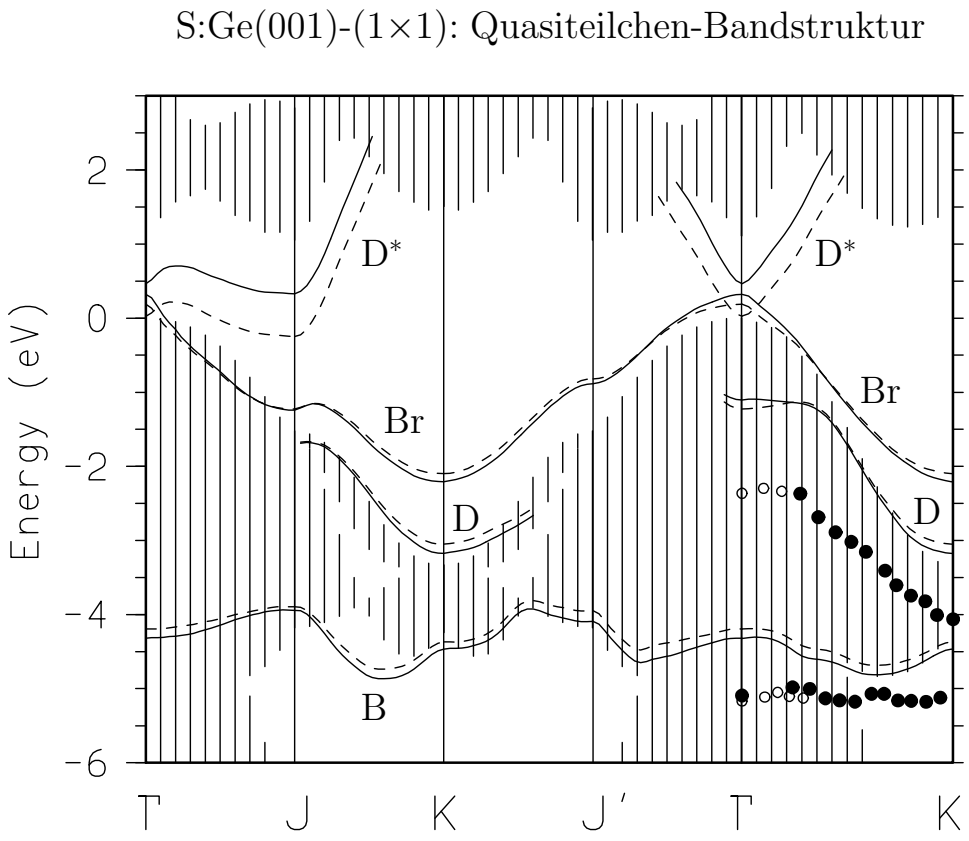


Abbildung 9.14: *Quasiteilchen-Bandstruktur der S:Ge(001)-(2×1) Oberfläche im Vergleich mit der LDA-Bandstruktur (gestrichelt) und mit experimentellen ARPES-Daten von Weser et al. [Wes88].*

einen metallischen Überlapp der Bänder Br und D* von 0.15 eV fest. Analog zur C-terminierten β-SiC(001) Oberfläche (vgl. Abschnitt 8.3) erwarten wir, daß dieser geringe Überlapp durch die Quasiteilchen-Korrekturen überwunden wird, denn die Quasiteilchen-Korrektur zum Volumengap von Ge ist mit 0.61 eV viermal größer. Wir gehen daher bei der Konstruktion des GW Selbstenergie-Operators analog zur β-SiC(001) Oberfläche vor und besetzten den Zustand Br am Γ-Punkt, obwohl er eine höhere LDA-Energie besitzt als der Zustand D*, den wir unbesetzt lassen. Diese Vorgehensweise stellt sich als richtig heraus, denn in der Quasiteilchen-Bandstruktur liegt die Energie von Br unter der des Zustands D*, und die beiden Zustände werden energetisch getrennt. Die Oberfläche ist also halbleitend mit einem direkten fundamentalen Quasiteilchen-Gap von 0.15 eV. Die Gap-Korrektur fällt mit +0.30 eV geringer aus als an der reinen Ge(001)-(2×1) Oberfläche (+0.40 eV). Der Grund dafür liegt in der unterschiedlichen Terminierung der beiden Oberflächen mit S bzw.

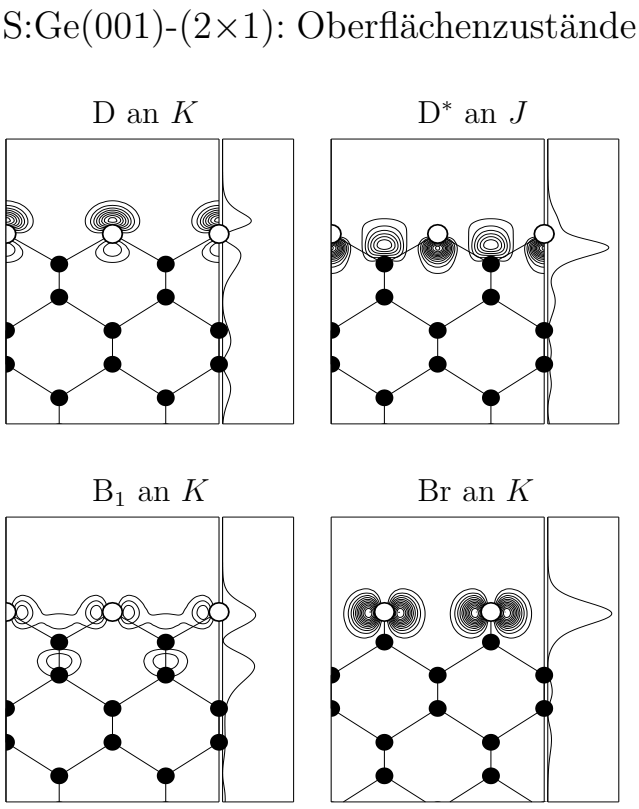


Abbildung 9.15: *Ladungsdichten der Oberflächenzustände an der S:Ge(001)-(2×1) Oberfläche (am K-Punkt bzw. J-Punkt).*

Ge. Aufgrund unserer Ergebnisse erwarten wir, daß die Fermi-Energie des Systems auf den kleinen Energiebereich zwischen 0.30 eV (höchster besetzter Oberflächenzustand) und 0.45 eV (tiefster unbesetzter Oberflächenzustand) über dem Volumen-VBM beschränkt ist. Weser et al. stellen eine Fermi-Energie von 0.31 eV über dem Volumen-VBM fest [Wes88], die innerhalb dieses Intervalls liegt. Weitere experimentelle Daten zur Gap-Energie oder zu den unbesetzten Zuständen sind uns nicht bekannt.

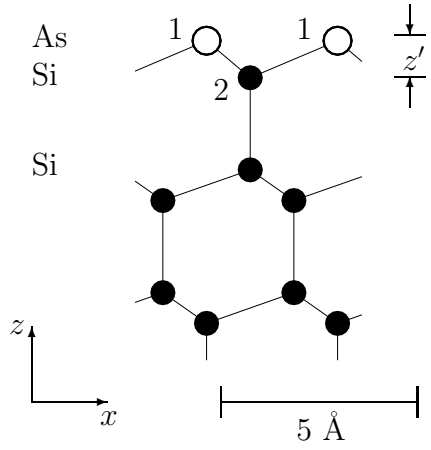
Weser et al. haben an der S:Ge(001)-(1×1) Oberfläche die elektronische Bandstruktur mittels ARPES-Untersuchungen bestimmt [Wes88]. Entlang der Hochsymmetrielinie ΓK stellen sie im Energiebereich von –6 eV bis 0 eV zwei Oberflächenbänder fest, die den von uns berechneten Bändern D und B zuzuordnen sind. Die Daten sind in Abb. 9.14 als einzelne Symbole dargestellt. Bezüglich der Dispersion der Bänder stellen wir für beide Zustände eine gute Übereinstimmung zwischen unseren berechneten Quasiteilchen-Energien und den experimentellen Daten fest. Die

Bandbreite des D Zustands zwischen Γ und K beträgt in unserer Quasiteilchen-Bandstruktur 2.05 eV in guter Übereinstimmung mit der gemessenen Bandbreite von 1.7 eV. Der B Zustand besitzt im Experiment praktisch keine Dispersion; er streut nur um etwa 0.2 eV. Wir erhalten für das B Band in unserer GW-Rechnung ebenfalls nur eine sehr geringe Dispersion von 0.5 eV auf der Linie ΓK. Die absolute energetische Lage der beiden Bänder in Bezug auf das Volumen-VBM ist hingegen in unserer Quasiteilchen-Bandstruktur deutlich höher als im Experiment. Der Unterschied beträgt 0.8 eV bis 1.2 eV für das D Band und 0.2 eV bis 0.7 eV für das B Band. Das gilt für die LDA-Bänder ebenso wie für die GWA-Bänder, denn die Quasiteilchen-Korrekturen sind sehr gering.

In [Krü90] war vermutet worden, daß die Diskrepanzen zwischen den LDA-Bändern und den experimentellen Daten auf systematische Fehler der LDA zurückzuführen seien und durch Quasiteilchen-Korrekturen behoben werden könnten, die zu einer Energieabsenkung der Bänder D und B und zu einer besseren Übereinstimmung mit dem Experiment führen. Diese Vermutung wird durch unsere Ergebnisse nicht bestätigt. Quasiteilchen-Korrekturen wurden für eine Vielzahl reiner und adsorbatbedeckter Halbleiter-Oberflächen berechnet (siehe die vorangegangen Kapitel sowie [Hyb87a, Hyb88b, Zhu89a, Nor91, Zhu91a, Hri93, Kre93, Nor93, Bla94, Kre95]). In den meisten Fällen sind die Korrekturen für besetzte Oberflächenzustände in der Nähe des fundamentalen Gaps viel kleiner als 0.5 eV. Die Oberflächenzustände an H-terminierten Oberflächen stellen hier eine Ausnahme dar (vgl. die Abschnitte 9.1 und 9.2 sowie [Bla94]). Aber im Falle der Schwefel-bedeckten S:Ge(001)-(1×1) Oberfläche sind Korrekturen der Bänder B und D in der Größenordnung 1 eV nicht zu erwarten und werden in unserer Rechnung auch nicht erzielt. Die Diskrepanz zwischen Theorie und Experiment bleibt auch in der GWA bestehen und wird durch Quasiteilchen-Korrekturen nicht überwunden.

Weser et al. geben auch Bandstrukturdaten für die ΓJ und $\Gamma J'$ Richtungen an [Wes88]. Allerdings beobachten die Autoren in ihren Spektren Photoemission von beiden Richtungen gleichzeitig, denn die verwendete Probe ist zweidomänig und liefert Überlagerungen von (2×1) und (1×2) Strukturen in den Spektren. Wir verzichten daher auf einen Vergleich unserer Ergebnisse mit den Daten. Abschließend merken wir an, daß der vollständig besetzte Br Zustand im Photoemissionsspektrum nicht beobachtet wird. Das hängt möglicherweise damit zusammen, daß der Zustand aus S 3*p* Orbitalen parallel zur Oberflächenebene besteht und somit nur eine sehr schwache Photoemissions-Intensität aufweist.

As:Si(111)-(1×1): Oberflächengeometrie



Atom	z' [Å]
1 (As)	0.935
	Δz [Å]
2 (Si)	0.000

Abbildung 9.16: *Geometrie der As:Si(111)-(1×1) Oberfläche [Hyb87a, Hyb88b]. Die offenen Kreise markieren As-Atome, die Punkte Si-Atome.*

9.4 Die As:Si(111)-(1×1) Oberfläche

Zum Abschluß dieses Kapitels wenden wir uns der Arsen-terminierten As:Si(111)-(1×1) Oberfläche zu. Arsen ist ein Element der fünften Hauptgruppe. Es besitzt fünf Valenzelektronen und bildet bevorzugt drei kovalente Bindungen. Die Anlagerung von As an der Si(111) Oberfläche geschieht daher an der in Abb. 9.9b vorgestellten Kristallebene, in der jedes Si Oberflächenatom drei freie Bindungen aufweist. Die As Adatome bilden jeweils drei chemische Si-As Bindungen aus. Sie erhalten dadurch eine Edelgaskonfiguration, und die Oberfläche wird passiviert. Nach der Adsorption einer ganzen Monolage As verhält sich die Oberfläche chemisch inert. Sie weist dann eine (1×1)-Rekonstruktion auf.

Die atomare Geometrie der Oberfläche wurde von Hybertsen und Louie [Hyb87a, Hyb88b] im Rahmen der LDA bestimmt. Die Relaxationen sind durchweg sehr gering. Das As Atom wird gegenüber seiner idealen Substitutions-Position in Richtung Vakuum verschoben, denn das As Atom weist als ein Element der vierten Periode eine größere Ausdehnung auf als Si, das der dritten Periode angehört. Die Relaxation beträgt 0.16 Å. Das bewirkt eine Verlängerung der Bindung zum benachbarten Si Atom auf 2.38 Å im Vergleich zur Bindungslänge von 2.33 Å im Si Volumenkristall. Alle anderen Atome erfahren praktisch keine Relaxation und verbleiben in ihren idealen Positionen. Wir stellen die Geometrie in Abb. 9.16 vor.

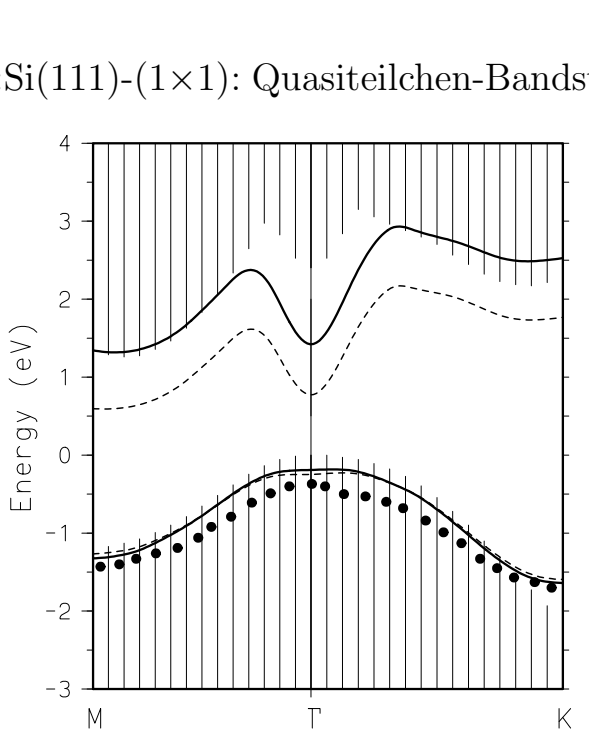


Abbildung 9.17: *GW Quasiteilchen-Bandstruktur der As:Si(111)-(1×1) Oberfläche im Vergleich mit der LDA Bandstruktur (gestrichelt) und mit experimentellen ARPES-Daten (Punkte) von Uhrberg et al. [Uhr87].*

In Abb. 9.17 vergleichen wir die zugehörige Quasiteilchen-Bandstruktur mit den zugrunde liegenden LDA-Energien (gestrichelt) und mit experimentellen ARPES-Daten von Uhrberg et al. [Uhr87]. Als Folge der As-Terminierung treten im Gap-Bereich zwei elektronische Oberflächenzustände auf. Hierin unterscheidet sich die Oberfläche deutlich von der in Abschnitt 9.2 diskutierten H:Si(111)-(1×1) Oberfläche. Bei letzterem System wird eine deutlich stärkere H–Si Bindung etabliert, die zu einer Absenkung der Oberflächenzustände um etwa 5 eV führt. Das Gap ist dementsprechend frei von Oberflächenzuständen. Hingegen ist im Falle der As-terminierten Oberfläche der Unterschied zwischen den Si-Si und den As-Si Bindungen viel geringer, so daß Zustände im Gapbereich nahe der projizierten Volumenbandstruktur resultieren. In Abb. 9.18 stellen wir die zugehörigen Dichteverteilungen (am M-Punkt) dar. Der energetisch niedrigere, besetzte Zustand wird im wesentlichen von den As $4p_z$ Orbitalen gebildet. Der andere, unbesetzte Zustand kommt durch die Hybridisierung von $4p$- und $4d$-Orbitalen der As Atome mit $2s$- und $2p$-Orbitalen der benachbarten Si Atome zustande. Die Bandstrukturenergien an den Hochsymmetriepunkten sind in Tab. 9.2 zusammengefaßt, zusammen

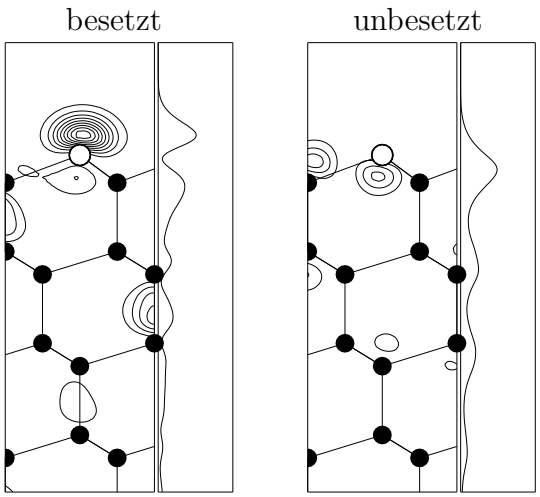


Abbildung 9.18: *Ladungsdichten der Oberflächenzustände an der As:Si(111)-(1×1) Oberfläche (am M-Punkt).*

mit Ergebnisse aus einer PW-GWA Rechnung von Hybertsen und Louie [Hyb87a, Hyb88b]. Der besetzte Zustand erfährt nur eine sehr geringe Energieabsenkung durch die Quasiteilchen-Korrekturen. Die von uns berechneten Quasiteilchen-Energien des Bandes liegen um 0.15 eV über den experimentellen Werten. Die Bandbreite wird geringfügig um 0.11 eV auf 1.45 eV vergrößert. Dieser Wert stimmt gut mit dem experimentellen Befund von 1.36 eV überein. Der unbesetzte Zustand wird um +0.65...+0.75 eV zu höheren Energien verschoben. Das direkte Oberflächen-Gap am Γ-Punkt wird dadurch um +0.60 eV vergrößert. Diese Quasiteilchen-Korrektur für das Oberflächen-Gap ist größer als an der reinen Si(001)-(2×1) Oberfläche (+0.50 eV) und an der reinen Si(111)-(2×1) Oberfläche (+0.35 eV) [Nor91]. Das liegt daran, daß die As-Terminierung die Eigenschaften des Selbstenergie-Operators an der Oberfläche entscheidend bestimmt und er sich von dem Selbstenergie-Operator an einer reinen Si Oberfläche unterscheidet. Unsere Quasiteilchen-Energien für den unbesetzten Oberflächenzustand stimmen gut mit den Ergebnissen von Hybertsen und Louie überein. Lediglich am Γ-Punkt erhalten die Autoren eine um 0.4 eV höhere Energie. Das liegt möglicherweise daran, daß im allgemeinen bereits in der LDA die Energien der Oberflächenzustände am Γ-Punkt eine größere Unsicherheit aufweisen als Zustände an anderen Punkten der SBZ [Hyb88b]. Wir erzielen für den Zustand am Γ-Punkt eine Quasiteilchen-Korrektur von +0.65 eV, die gut mit der

As:Si(111)-(1×1): Quasiteilchen-Energien

[eV]	LDA	GWA	GWA [Hyb88b]	Exp. [Uhr87]
M	–1.27	–1.32	–1.44	–1.44
	0.60	1.34	1.45	
Γ	–0.25	–0.19	–0.39	–0.39
	0.77	1.42	1.80	
K	–1.59	–1.64	–1.71	–1.75
	1.77	2.53	2.45	

Tabelle 9.2: *Bandstruktur-Energien für die Oberflächenzustände an der As:Si(111)-(1×1) Oberfläche (in eV). Erste Spalte: LDA. Zweite Spalte: GO-GWA, basierend auf der in RPA berechneten statischen dielektrischen Matrix. Dritte Spalte: analoge GW-Rechnung, jedoch mit ebenen Wellen (PW-GWA) [Hyb88b]. Vierte Spalte: winkelaufgelöste Photoemissionsdaten (ARPES) von Uhrberg et al. [Uhr87].*

von Hybertsen und Louie erzielten Korrektur von +0.8 eV übereinstimmt. In einer Tunnelspektroskopie-Messung von Uhrberg wurde ein Tunnelgap zwischen 1.9 eV und 2.2 eV beobachtet [Uhr87]. Unser Ergebnis für das direkte Quasiteilchen-Oberflächengap am Γ-Punkt beträgt 1.61 eV. Es ist etwas kleiner als der experimentelle Wert.

Kapitel 10

Ionisierungsenergien

Wir haben bislang alle Bandstrukturenergien auf das Volumen-Valenzbandmaximum (VBM) bezogen, das wir als Referenzenergie mit dem Wert 0 eV bezeichnet hatten. Das Studium eines Oberflächensystems eröffnet jedoch auch die Möglichkeit, die Bandstrukturenergien direkt mit dem Vakuum-Energieniveau zu vergleichen (siehe hierzu Abb. 6.3). Von besonderem Interesse ist hierbei die Energiedifferenz zwischen dem höchsten besetzten Zustand und dem Vakuum-Niveau. Diese Differenz definiert die Ionisierungsenergie der Oberfläche und ist somit experimentell zugänglich [Mön93]. Bei der Ionisierung handelt es sich um einen Vorgang, bei dem ein Elektron vollständig aus dem System entfernt wird ($N \rightarrow N-1$). Daher stellt die Quasiteilchen-Bandstruktur auch für diese Größe den konsistenten Anknüpfungspunkt dar, während die LDA unter systematischen Verfälschungen leidet.

Wir diskutieren im folgenden für verschiedene Oberflächen die Ionisierungsenergien, die sich aus unseren Rechnungen ergeben. Vorweg merken wir die folgenden Punkte an.

- Es ist unklar, inwieweit Oberflächenzustände zur Ionisierungsenergie beitragen, oder ob die Ionisierungsenergie primär einen Volumeneffekt darstellt. Wir nehmen in dieser Arbeit an, daß die Ionisierung im wesentlichen durch Übergänge von Volumenzuständen ins Vakuum zustande kommt, und geben daher die Differenz des Volumen-VBM zum Vakuumniveau als Ionisierungsenergie an.

- Falls in einer Slabrechnung die beiden Slaboberflächen unterschiedlich terminiert sind, so kann es zu Aufladungen der Oberfläche und zu Verfälschungen der Ionisierungsenergie kommen. Außerdem kann dann nicht zwischen den Ioni-

sierungsenergien der beiden verschiedenen Oberflächen unterschieden werden, denn das System gibt für beide Oberflächen denselben Wert an. Um diese Probleme zu umgehen, werden für die Rechnungen in diesem Kapitel beide Oberflächen eines jeden Slabs gleichartig aufgebaut.

- Durch die Quasiteilchen-Korrekturen erfährt das Volumen-VBM eine Veränderung; es wird bei den hier untersuchten Materialien im Vergleich zur LDA abgesenkt. Da das Vakuumniveau durch die Quasiteilchen-Korrekturen nicht verändert wird, erfährt die Ionisierungsenergie durch die Quasiteilchen-Korrekturen eine entsprechende Anhebung gegenüber dem LDA-Resultat [Zhu91a]. Die GWA-Ergebnisse sollten allerdings nur als vorläufiges Resultat für die echte Quasiteilchen-Ionisierungsenergie verstanden werden. Die GWA stellt nur eine recht einfache Näherung für den exakten Selbstenergie-Operator dar. Die hervorragende Übereinstimmung der GW Quasiteilchen-Bandstrukturen mit experimentellen Resultaten zeigt, daß Energie*differenzen* zwischen verschiedenen Zuständen mit sehr hoher Genauigkeit bestimmt werden können. Ob das auch für absolute Energien in Bezug auf das Vakuumniveau zutrifft, ist unklar. Die hier vorgestellten Ergebnisse sollten daher nicht überinterpretiert werden.

- Die Ionisierungsenergie muß von der Austrittsarbeit unterschieden werden. Letztere ist durch die Energiedifferenz der Fermienergie zum Vakuumniveau gegeben. Sie unterscheidet sich also von der Ionisierungsenergie durch die Differenz zwischen Fermienergie und VBM [Mön93].

Kommen wir nun zu den einzelnen Oberflächen. In Tabelle 10.1 sind alle Ergebnisse für die Ionisierungsenergien der Oberflächen zusammengefaßt, die wir in den vergangenen Kapiteln untersucht haben. Ausgenommen sind das SDM für die Si(001)-(2×1) Oberfläche, da sie keine realistische Geometrie darstellt, sowie die β-SiC(001)-(2×1) Oberfläche. Um letzere mit einem beidseitig C-terminierten Slab mit Dimergeometrie zu behandeln, bedarf es einer zweifach größeren $p(2\times2)$ Einheitszelle. Der dafür notwendige Aufwand erscheint uns zu groß. Die Vergrößerung der Ionisierungsenergie durch die Quasiteilchen-Korrekturen hängt von der Terminierung der jeweiligen Oberfläche nicht ab, sondern ist durch das Substrat gegeben. Diese Korrektur beträgt +0.41 eV für Si und +0.22 eV für Ge. Für β-SiC geben wir den Wert der Vollständigkeit halber auch an; er beträgt +0.69 eV.

Unser Ergebnis für die Quasiteilchen-Ionisierungsenergie für die Si(001)-(2×1)

Oberflächen-Ionisierungsenergien

[eV]	LDA	GWA	Exp.
Si(001)-(2×1)	4.71	5.12	5.33[a]
H:Si(111)-(1×1)	4.50	4.91	6.6[a]
As:Si(111)-(1×1)	4.38	4.79	
Ge(001)-(2×1)	4.12	4.34	4.5[b]
H:Ge(001)-(2×1)	3.79	4.01	
S:Ge(001)-(1×1)	5.32	5.54	

[a]Ref. [Mön93] [b]Ref. [Kip96]

Tabelle 10.1: *Berechnete Ionisierungsenergien verschiedener Oberflächen von Si und Ge.*

Oberfläche stimmt sehr gut mit dem gemessenen Wert überein, während das LDA-Ergebnis zu klein ist. Im Falle der H-bedeckten H:Si(111)-(1×1) Oberfläche liegen die berechneten Werte hingegen um 2.1 eV (LDA) bzw. 1.7 eV (GWA) unter dem experimentellen Befund.

Im Falle der reinen Ge(001)-(2×1) Oberfläche stimmt unser GW-Resultat hervorragend mit der gemessenen Ionisierungsenergie [Kip96] überein. Durch die Adsorption einer Monolage Wasserstoff wird die Ionisierungsenergie um 0.33 eV abgesenkt. Dies entspricht dem experimentellen Befund für die ähnlich strukturierte Si(001) Oberfläche, deren Ionisierungsenergie durch die Adsorption von Wasserstoff ebenfalls um bis zu 0.42 eV reduziert wird [Mön93]. Für die Adsorption von Schwefel auf der Ge(001) Oberfläche stellen wir hingegen eine Anhebung der Ionisierungsenergie um 1.20 eV fest. Die große Differenz der Elektronegativitäten von Ge (2.0) und S (2.6) führt zu elektronischem Ladungstransport zu den Adatomen und somit zur Bildung elektrostatischer Dipole an der Oberfläche. Dadurch wird die Potentialschwelle zwischen Substrat und Vakuum vergrößert und die Ionisierungsenergie heraufgesetzt.

Teil III

Semicore- und Core-Elektronen in Halbleitern

Kapitel 11

Vorbemerkungen

11.1 Lokalisierte Zustände

In diesem Abschnitt unserer Arbeit wenden wir uns elektronischen Zuständen zu, die durch eine starke Lokalisierung auf kleinem Raum geprägt sind. Hierzu müssen wir zunächst präzisieren, was wir unter Lokalisierung verstehen. Insbesondere müssen wir eine Abgrenzung zum Begriff der Lokalisierung an Oberflächen, die wir im letzten Abschnitt betrachtet hatten, herstellen.
(i) An Oberflächen, Defekten und bei Quantenstrukturen bilden sich infolge der auf atomarer Bindungslängenskala gestörten Kristallperiodizität Zustände aus, deren Einteilchen-Wellenfunktionen außerhalb des Bereichs der Störung exponentiell abklingen. Die Zustände sind auf diesen Bereich eingeschränkt bzw. an der Störung der Periodizität lokalisiert. Diese Lokalisierung kann in einer, zwei oder drei Dimensionen auftreten. Die Behandlung solcher lokalisierter Zustände ist gegenüber einem einfachen Volumenkristall mit einem erhöhten numerischen Aufwand verbunden. In jeder Raumrichtung, bezüglich derer die Lokalisierung auftritt, muß die Einheitszelle soweit verlängert werden, daß der lokalisierte Zustand vollständig hineinpaßt. Die Oberflächen-Systeme, die wir im letzten Abschnitt behandelt haben, zählen zu dieser Klasse von Lokalisierungs-Problematik in einer Dimension. Paradoxerweise ist die Behandlung dieses Typs von Lokalisierung um so aufwendiger, je größer der Lokalisierungsbereich ist, d.h. je weniger ausgeprägt die Lokalisierung erscheint. Relevant ist hier das Verhältnis zwischen der Ausdehnung des lokalisierten Zustands und der atomaren Bindungslänge.
(ii) Ein ganz anderer Problemtyp eröffnet sich bei elektronischen Zuständen, deren

räumliche Ausdehnung wesentlich kleiner ist als die Bindungslängenskala. In einem Atom finden sich einerseits räumlich weiter ausgedehnte Valenzelektronen, die im Kristall zur chemischen Bindung führen und in die delokalisierten Valenzzustände übergehen. Andererseits enthält das Atom auch stärker lokalisierte Core-Elektronen. Diese sind durch wesentlich steilere Wellenfunktionen und eine Beschränkung auf einen sehr kleinen Raumbereich gekennzeichnet. Wenn das Atom die Kernladungszahl Z besitzt, so nimmt die am stärksten lokalisierte $1s$ Wellenfunktion schon bei einem Radius von etwa a_0/Z das Maximum ihrer Ladungsdichte an, also auf einer Längenskala, die ein oder zwei Größenordnungen kleiner ist als die zwischenatomare Bindungslängenskala. Als relevante Größe zur Beschreibung dieses Typs von Lokalisierung können wir die Position von Wellenfunktions-Maxima und -Knoten bzw. die Steilheit von Wellenfunktionen heranziehen. Wenn die Festkörpereigenschaften eines Systems untersucht werden sollen, so gibt es dazu zwei mögliche Vorgehensweisen. Entweder verwendet man Basisfunktionen, die für alle relevanten Längenskalen gleich gut geeignet sind, oder man muß die inneren Elektronen eliminieren. Der zweite Weg führt auf das Pseudopotentialkonzept. Wir werden in dieser Arbeit aufzeigen, welche Probleme und systematischen Fehler beim Pseudopotentialkonzept auftreten können.

Bei der Behandlung der H-terminierten Halbleiteroberflächen in Kap. 9 traten beide Typen von Lokalisierung gleichzeitig auf. Einerseits sind die H-abgeleiteten Oberflächenzustände wegen der gestörten Periodizität an den Oberflächenatomen der ersten und zweiten Schicht lokalisiert. Andererseits weisen die zugehörigen Wellenfunktionen an den H-Adatomen die typische Steilheit des H $1s$ Orbitals auf.

Wir werden hauptsächlich zwei Fragestellungen untersuchen, die im Zusammenhang mit den inneren Elektronen auftreten.

- Um zu einer möglichst einfachen Beschreibung von strukturellen und elektronischen Kristalleigenschaften zu gelangen, wird man zunächst nur die äußersten Valenzelektronen im Pseudopotential berücksichtigen. Wir werden in Kap. 13 Fälle untersuchen, in denen auch die Zustände der nächsten Schale als Valenzzustände implementiert werden müssen, um zu konvergenten Ergebnissen für die Valenz- und Leitungsbandstruktur zu gelangen, die dann mit experimentellen Daten gut übereinstimmen. Das ist zuvorderst bei den II-VI Halbleitern der Fall. Hier untersuchen wir als Prototyp die kubische Modifikation von CdS.

- In den letzten Kapiteln der Arbeit werden die Einteilchenenergien der lo-

kalisierten inneren Elektronen selbst im Vordergrund stehen. Schon bei der Behandlung von H-terminierten Halbleiteroberflächen hatten wir festgestellt, daß die LDA die Energien der H-abgeleiteten Oberflächenzustände, die steile Wellenfunktionen aufweisen, systematisch zu hoch schätzt. Bei den wesentlich stärker lokalisierten Core-Elektronen tritt diese Problematik noch weit ausgeprägter in Erscheinung. Wir werden in Kap. 14 mit Hilfe eines erweiterten GWA Formalismus Quasiteilchen-Energien für solche Zustände berechnen, die mit den experimentellen Daten wesentlich besser übereinstimmen.

Wir werden all diese Fragestellungen in folgender Weise behandeln. Wir konstruieren zunächst im Rahmen der LDA Pseudopotentiale, die alle zu behandelnden elektronischen Zustände in Form von Valenzzuständen beschreiben. Mit Hilfe dieser LDA-Pseudopotentiale führen wir LDA-Rechnungen für die zu untersuchenden Festkörper durch. Schließlich konstruieren wir auf der Basis der LDA-Ergebnisse den GW-Selbstenergie-Operator und beschreiben mit ihm die Austausch-Korrelationseffekte der Elektronen. Das Austausch-Korrelations-Potential der LDA wird dabei eliminiert. Die GW-Ergebnisse für die Bandstruktur-Energien können dann direkt mit experimentellen Daten verglichen werden.

11.2 Selbstwechselwirkungen

Gerade bei der Untersuchung stark lokalisierter Zustände stellt sich die Frage nach der Behandlung von Selbstwechselwirkungen. [Dieser Begriff hat nichts mit dem Begriff der Selbst*energie* zu tun.] Dabei handelt es sich um ein Artefakt, der bei vielen effektiven Einteilchen-Zugängen zum Vielelektronenproblem auftritt. Im exakten Vielteilchen-Hamiltonian gibt es keine Wechselwirkung eines Elektrons mit sich selbst. Bei der Verwendung einer Einteilchen-Methode wird ein effektiver Einteilchen-Hamiltonian konstruiert, der die Elektron-Elektron Wechselwirkung in vereinfachter Weise beinhaltet. Ein typisches Beispiel sind das Coulomb- und das Austausch-Korrelations-Potential in der Dichtefunktionaltheorie. In diese Größen geht die elektronische Einteilchen-Dichte ein. Im Rahmen des Modellhamiltonians kann dann ein Elektron über diese Dichte, in die es selber eingeht, mit sich selbst wechselwirken. Diese unphysikalische Selbst-Wechselwirkung kann in vielen Situationen zu einer systematischen Verfälschung von Bandstruktur-Ergebnissen führen.

Es wurde häufig gemutmaßt, daß die GW-Näherung frei von Selbstwechselwir-

kungen sei. Unserer Meinung nach ist das falsch. Wir betrachten als einfachstes Beispiel das Wasserstoff-Atom. Hier liegt eine besonders einfache Situation vor, denn das System enthält nur ein Elektron, und es tritt gar keine Elektron-Elektron Wechselwirkung auf. Im Rahmen der LDA errechnet man aus der Wellenfunktion des einen Elektrons eine Ladungsdichte und konstruiert daraus das Coulomb-Potential und das Austausch-Korrelations-Potential. Die beiden Potentiale heben sich nicht auf, und ihre Summe stellt eine Selbstwechselwirkung dar. Im Rahmen der GWA tritt zusätzlich zum Coulomb-Potential der Selbstenergie-Operator auf. Er besteht aus dem Austausch-Anteil und dem Korrelations-Anteil. Der Austausch-Anteil und das Coulomb-Potential heben sich im Falle eines einzelnen Elektrons exakt auf. Es kann leicht gezeigt werden, daß sich die Selbstwechselwirkungsterme des Coulomb-Potentials und des Austauschenergie-Operators in jedem Elektronensystem gegenseitig aufheben, so daß die Summe beider Terme selbstwechselwirkungsfrei ist. In der Hartree-Fock Theorie treten nur diese beiden Terme der Elektron-Elektron-Wechselwirkung auf; dementsprechend gibt es in der Hartree-Fock Theorie keine Selbstwechselwirkung. Der Korrelations-Anteil des Selbstenergie-Operators, der in der GWA hinzukommt, ist aber im Fall des H-Atoms sicher nicht exakt Null. Das Wasserstoff-Atom besitzt eine – wenn auch geringe – Polarisierbarkeit (2.7). Sie führt letztlich zu Abschirmeffekten im H-Atom und im Rahmen der GWA zu einer unphysikalischen „Selbst-Korrelation“ des Elektrons mit sich selbst. Erst der exakte Selbstenergie-Operator, der über die GWA hinaus sämtliche Vertex-Korrekturen berücksichtigt, wird wieder frei von Selbstwechselwirkungen sein.

Die Bandstruktur-Energien stärker lokalisierter Zustände werden von der LDA grundsätzlich zu hoch eingeschätzt. Ein Beispiel hierfür sind die kationischen Semicore *d*-Zustände in II-VI Halbleitern (siehe Kap. 13). Für diese systematischen Abweichungen ist zum Teil die Selbstwechselwirkung der Zustände verantwortlich [Vog95]. Sie läßt sich innerhalb der LDA durch entsprechende Selbstwechselwirkungs-Korrekturen (SIC) teilweise kompensieren [Per81, Sva90]. Diese Technik wird als SIC-LDA bezeichnet. In vielen Fällen wird die Selbstwechselwirkung eines lokalisierten Zustands im Atom und im Festkörper ähnlich ausfallen. Unter dieser Voraussetzung kann die Korrektur bereits im Atom bei der Konstruktion der Pseudopotentiale vorgenommen werden [Vog95, Vog96]. Die Anwendung dieser korrigierten SIC-Pseudopotentiale auf den Kristall verläuft dann völlig analog zur gewöhnlichen LDA-Technik. Die *d*-Zustände in II-VI Halbleitern erfahren durch die Korrektur eine erhebliche Energieabsenkung [Vog95, Vog96].

Im Rahmen der GWA ist es wesentlich schwieriger, über die prinzipielle Existenz von Selbstwechselwirkungen hinaus quantitative Aussagen über ihre Größe zu treffen. Es wurde daher oft der Vorschlag gemacht, die GWA mit der in LDA bestimmten Selbstwechselwirkungs-Korrektur ($\mathrm{SIC_{LDA}}$) zu kombinieren. Das könnte etwa geschehen, indem die GWA auf der Basis von SIC-LDA-Pseudopotentialen durchgeführt wird. Man würde daraus GW-Quasiteilchen-Ergebnisse erhalten, die mittels $\mathrm{SIC_{LDA}}$ korrigiert sind. Eine mit der GWA konsistente Selbstwechselwirkungs-Korrektur muß aber mit der entsprechenden SIC-Korrektur der GWA ($\mathrm{SIC_{GWA}}$) geschehen. Diese Korrektur $\mathrm{SIC_{GWA}}$ ist jedoch äußerst schwer zu behandeln. Wenn man die Korrektur statt dessen mit $\mathrm{SIC_{LDA}}$ durchführt, so suggeriert man mit der Bezeichnung „selbstwechselwirkungsfrei“ eine systematische Verbesserung über die Standard-GWA hinaus. Die Differenz zwischen $\mathrm{SIC_{GWA}}$ und $\mathrm{SIC_{LDA}}$ stellt aber eine nach heutigem Kenntnisstand noch nicht zu kontrollierende Größe dar, von der wir noch nicht einmal sagen können, welches Vorzeichen sie besitzt und ob sie groß oder klein ist. Daher sind GW-Ergebnisse auf der Grundlage von SIC-LDA-Pseudopotentialen nicht verläßlicher als die der Standard-GWA, selbst wenn sie – aus welchen Gründen auch immer – besser mit experimentellen Daten übereinstimmen sollten. Sie bergen im Gegenteil die Gefahr in sich, daß man ihre Aussagekraft überschätzt. Wir verzichten daher in dieser Arbeit auf jede Art von Selbstwechselwirkungs-Korrektur und behandeln alle Probleme im Rahmen der Standard-GWA bzw. mit systematischen Erweiterungen innerhalb der GWA.

Kapitel 12

Technische Vorbemerkungen

Die Behandlung der Core- und Semicore-Zustände erfordert einige Modifikationen des bisher verfolgten Standard-Zugangs. Das betrifft vor allem die Realisierung der LDA und die Konstruktion geeigneter Pseudopotentiale.

12.1 Durchführung der LDA-Rechnungen

In unseren LDA-Rechnungen werden zur Darstellung der Wellenfunktionen lokalisierte Gauß-Orbitale verwendet. Die Ladungsdichte $\rho(\mathbf{r})$ sowie die lokalen Potentiale $V(\mathbf{r})$ haben wir hingegen bisher mit ebenen Wellen (PW) ausgedrückt. Wenn im System stark lokalisierte Zustände vorliegen, so erweist sich diese PW-Darstellung als sehr aufwendig. Besonders bei zwei Rechenschritten führt sie zu extrem langen Rechenzeiten innerhalb der LDA:
(1) Bei der Bestimmung der Hamilton-Matrix müssen Matrixelemente $\langle\alpha|V|\alpha'\rangle$ des lokalen Potentials V mit den Basisfunktionen $|\alpha\rangle$ der Wellenfunktionen berechnet werden. Dabei tritt eine Summe über die ebenen Wellen auf.
(2) Zur PW-Basis gehört ein äquivalentes Raumnetz, dessen Mächtigkeit der des PW-Basissatzes entspricht. Für jeden dieser Raumpunkte muß die Ladungsdichte berechnet werden.
Der numerische Aufwand beider Programmpunkte steigt linear mit der Größe des PW-Basissatzes. In vielen Fällen erweist es sich als äußerst zeitsparend, wenn die Ladungsdichte und das Potential ebenfalls in eine Gauß-Basis transformiert werden, die aus wesentlich weniger Ansatzfunktionen als die PW-Basis bestehen kann:

$$f(\mathbf{r}) = \sum_{\gamma} \langle f \rangle_{\gamma} \chi_{\gamma}(\mathbf{0}, \mathbf{r}) \quad . \tag{12.1}$$

Zu einer genaueren Darstellung dieser Transformation verweisen wir auf Anhang A.

Im folgenden treten vier Basissätze auf. Die Wellenfunktionen werden in einer Gauß-Basis $\{\alpha\}$ dargestellt. Die Darstellung der Felder in der *großen* PW-Basis bezeichnen wir mit $\{\mathbf{G}\}$ (dieser riesige Basissatz wird nach wie vor zur Speicherung der Felder verwendet – das führt aber höchstens zu Speicherproblemen, nicht zu Rechenzeit-Problemen. Alle noch verbleibenden Operationen in dieser großen PW-Basis sind sehr schnell – wie etwa Fouriertransformationen). Die zusätzliche Gauß-Basis, in die wir die Felder umrechnen, bezeichnen wir mit $\{\gamma\}$. Bei der Bestimmung der Hamilton-Matrix erweist es sich als günstig, zusätzlich zu den Gauß-Funktionen eine kleine PW-Basis $\{\mathbf{g}\}$ zu verwenden. Sie sollte so gewählt werden, daß das zugehörige Raumnetz eine Untermenge des dichteren Raumnetzes zu $\{\mathbf{G}\}$ ist. Der Basissatz $\{\mathbf{g}\}$ unterscheidet sich von $\{\mathbf{G}\}$ durch seine wesentlich kleinere Abschneide-Energie. Wir wählen für $\{\mathbf{g}\}$ im allgemeinen um zwei Größenordnungen weniger ebene Wellen als für $\{\mathbf{G}\}$.

(Zu 1) Bei der Bestimmung der Hamilton-Matrix gehen wir folgendermaßen vor. Das lokale Potential liege in Form ebener Wellen vor:

$$V(\mathbf{r}) = \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}} \quad . \tag{12.2}$$

Die Umrechnung auf die Gauß-Basis gemäß Anhang A ergibt eine Darstellung

$$V^{\text{Gauß}}(\mathbf{r}) = \sum_{\gamma} \langle V \rangle_{\gamma} \chi_{\gamma}(\mathbf{0}, \mathbf{r}) \quad . \tag{12.3}$$

Analog zur Wellenfunktions-Basis $\{\alpha\}$ konzipieren wir die Gauß-Basis $\{\gamma\}$, indem wir die Gauß-Funktionen an den Atompositionen zentrieren (das ermöglicht besonders gut die Darstellung der steilen Radialanteile des Potentials um die Atomkerne herum). Im zwischenatomaren Bereich und im Vakuum eines Slabsystems klingen die Gauß-Orbitale auf Null ab (es sei denn, man berücksichtigt sehr kleine Abklingkonstanten γ oder plaziert weitere Gauß-Orbitale im Zwischenbereich – beides wäre sehr umständlich). Das Potential konvergiert aber im zwischenatomaren Bereich und im Vakuum eines Slabsystems aus rechentechnischen Gründen im allgemeinen nicht gegen Null, sondern gegen einen konstanten Wert. Um auch in diesen Bereichen das dort nur schwach veränderliche Potential korrekt zu erfassen, stellen wir die Abweichung zwischen der Gauß-Darstellung und der ursprünglichen Fourier-Darstellung mit der kleinen PW-Basis $\{\mathbf{g}\}$ dar. Die Differenz lautet im reellen Raum

$$\Delta V(\mathbf{r}) \;=\; V(\mathbf{r}) - V^{\text{Gauß}}(\mathbf{r})$$

$$= \sum_{\mathbf{G}} V_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}} - \sum_{\gamma} \langle V \rangle_{\gamma} \chi_{\gamma}(\mathbf{0}, \mathbf{r})$$

$$= \sum_{\mathbf{G}} \left(V_{\mathbf{G}} - \sum_{\gamma} \langle V \rangle_{\gamma} K_{\gamma}^{\mathbf{G}}(\mathbf{0}) \right) e^{i\mathbf{G}\mathbf{r}} \quad . \tag{12.4}$$

[$K_{\gamma}^{\mathbf{G}}(\mathbf{0})$ ist die Fouriertransformierte der Gauß-Funktion γ (siehe Anhang A).] Von den in der letzten Zeile bestimmten Fourierkomponenten der Differenz ΔV berücksichtigen wir nur diejenigen, die auch in der kleinen PW-Basis $\{\mathbf{g}\}$ enthalten sind ($\mathbf{G} \in \{\mathbf{g}\}$), und erhalten damit eine äußerst zuverlässige Darstellung der schwach veränderlichen Potentialdifferenz ΔV. Das gesamte Potential lautet nunmehr:

$$V(\mathbf{r}) = V^{\text{Gauß}}(\mathbf{r}) + \Delta V(\mathbf{r})$$

$$\doteq \sum_{\gamma} \langle V \rangle_{\gamma} \chi_{\gamma}(\mathbf{0}, \mathbf{r}) + \sum_{\mathbf{g}} \left(V_{\mathbf{g}} - \sum_{\gamma} \langle V \rangle_{\gamma} K_{\gamma}^{\mathbf{g}}(\mathbf{0}) \right) e^{i\mathbf{g}\mathbf{r}} \quad . \tag{12.5}$$

Die Berechnung der Matrixelemente des Potentials, die in den Hamiltonian eingehen, erfolgt nun gemäß

$$\langle \alpha | V | \alpha' \rangle = \sum_{\gamma} \langle V \rangle_{\gamma} \langle \alpha | \gamma | \alpha' \rangle + \sum_{\mathbf{g}} \left(V_{\mathbf{g}} - \sum_{\gamma} \langle V \rangle_{\gamma} K_{\gamma}^{\mathbf{g}}(\mathbf{0}) \right) \langle \alpha | e^{i\mathbf{g}\mathbf{r}} | \alpha' \rangle \quad . \tag{12.6}$$

(Zu 2) Die Berechnung der Ladungsdichte auf dem dichten Raumnetz, das zu der großen PW-Basis $\{\mathbf{G}\}$ gehört, erfolgte bisher in der Form

$$\rho(\mathbf{r}) = \sum_{\mathrm{n}\mathbf{k}} |\psi_{\mathrm{n}\mathbf{k}}(\mathbf{r})|^2 \tag{12.7}$$

$$\text{mit} \quad \psi_{\mathrm{n}\mathbf{k}}(\mathbf{r}) = \sum_{\alpha} C_{\mathrm{n}\mathbf{k}}^{\alpha} \chi_{\alpha}(\mathbf{k}, \mathbf{r}) \quad .$$

Diese Ortsdarstellung zeichnet sich besonders dadurch aus, daß nur in ihr die Ladungsdichte in eine Summe $\sum_{\mathrm{n}\mathbf{k}}$ zerfällt, in der jeder Term seinerseits ein einfaches Betragsquadrat $|\psi|^2$ einer ortsabhängigen Funktion ist. Diesem numerischen Vorteil steht bei lokalisierten Zuständen der entscheidende Nachteil gegenüber, den Ausdruck für eine ungeheuer große Zahl von Raumpunkten explizit berechnen zu müssen. In diesem Falle ergibt sich eine wesentliche schnellere Vorgehensweise, wenn die Ladungsdichte direkt in der Gauß-Basis bestimmt wird:

$$[\rho]_{\gamma} = \int \chi_{\gamma}^{*}(\mathbf{0}, \mathbf{r}) \rho(\mathbf{r}) d^3 r = \sum_{\mathrm{n}\mathbf{k}} \sum_{\alpha\alpha'} \left(C_{\mathrm{n}\mathbf{k}}^{\alpha} \right)^{*} C_{\mathrm{n}\mathbf{k}}^{\alpha'} \langle \alpha | \gamma | \alpha' \rangle \quad . \tag{12.8}$$

Die Struktur dieser Gleichung ist wegen der doppelten Summe über α und α' komplizierter als die direkte Berechnung der Ladungsdichte (12.7) für jeden Raumpunkt.

Wenn bei einem Festkörper nur die typischen Valenzelektronen explizit behandelt werden, so ist die Verwendung einer Gauß-Basis für die Ladungsdichte wesentlich aufwendiger als die direkte Berechnung. Bei der Untersuchung stark lokalisierter Semicore- und Core-Zustände gewinnt man hingegen durch die Gauß-Darstellung eine Größenordnung oder mehr an Rechenzeit. Die Ladungsdichte wird aus dieser Basis $\{\gamma\}$ anschließend wieder in die große PW-Basis $\{\mathbf{G}\}$ umgerechnet, um das Coulomb-Potential bestimmen zu können, und anschließend in der reellen Raum fouriertransformiert, um das Austausch-Korrelations-Potential zu berechnen.

Der Gewinn an Rechenzeit durch die Verwendung der Gauß-Orbitalbasis $\{\gamma\}$ beträgt bei einigen Systemen, wie z. B. bei Si bei Verwendung eines Si^{12+} PP, eine Größenordnung oder mehr. In einer Si^{12+} PP Rechnung benötigt man pro Si Atom etwa 10^5 ebene Wellen zur Darstellung der Ladungsdichte und des Potentials – oder eben nur 70 Gauß-Orbitale. Wir haben das Verfahren an verschiedenen Volumenkristallen getestet und die Ergebnisse mit denen konventioneller LDA-Rechnungen verglichen, in denen alle Felder nur mit ebenen Wellen dargestellt wurden. Die Bestimmung der totalen Energie mittels der Gauß-Darstellung $\{\gamma\}$ für die Felder unterliegt gewissen Ungenauigkeiten, die durch die Verwendung der relativ kleinen Basis zustande kommen. Zur Berechnung von Gitterkonstanten oder von Relaxationen ist die Vorgehensweise in der vorliegenden Form daher nur bedingt geeignet. Bei den Bandstrukturen stellen wir hingegen hervorragende Übereinstimmungen fest. Die Abweichungen sind in allen untersuchten Fällen kleiner als 0.05 eV. Die elektronische Struktur wird in unserem Zugang also sehr genau erfaßt, und die Methode ist zur Untersuchung der Energien von Semicore- und Core-Elektronen sowie ihres Einflusses auf die Bandstruktur ausgezeichnet geeignet – besonders in solchen Fällen, wo die konventionelle LDA wegen zu langer Rechenzeiten ausscheidet. Die Untersuchung von Core-Niveau Verschiebungen in Kap. 15 könnten wir ohne diesen Formalismus gar nicht durchführen.

12.2 Das Pseudopotentialkonzept

Von den Elektronen innerhalb eines Atoms sind vor allem die Valenzelektronen in der äußersten Schale für die Ausbildung chemischer Bindungen und für die charakteristischen Festkörpereigenschaften eines Kristalls verantwortlich. Die tieferliegenden Core-Elektronen scheinen auf den ersten Blick nichts zu den Festkörpereigenschaften beizutragen. Ihre Bandstruktur-Energien weisen meist nur eine sehr schwache Di-

spersion auf. Zwischen den Core-Elektronen an benachbarten Atomen besteht somit praktisch keine Wechselwirkung. Um zu einer Beschreibung der Festkörpereigenschaften zu kommen, sollte es im allgemeinen ausreichen, sich auf die Valenzelektronen zu beschränken und die energetisch tieferliegenden Elektronen aus dem Problem zu eliminieren. Von einer solchen Vorgehensweise kann man sich aus zwei Gründen eine erhebliche Reduktion des numerischen Aufwandes erhoffen. Zum einen reduziert sich die Anzahl der zu berücksichtigenden Elektronen. Zum anderen erfordert gerade die Darstellung der sehr steilen Wellenfunktionen der inneren Elektronen einen großen Aufwand an Basisfunktionen, der durch die Elimination dieser Zustände entfällt.

Die Eliminierung der inneren Elektronen erfolgt durch die Konstruktion eines Pseudopotentials für das zu beschreibende Atom. Für die verbleibenden Valenzelektronen ergibt sich dabei ebenfalls ein numerischer Vorteil. Im all-electron Problem müssen die Wellenfunktionen der Valenzelektronen orthogonal zu denen der inneren Elektronen sein. Das erzwingt Oszillationen im Radialanteil ihrer Wellenfunktion mit mehreren Nulldurchgängen. Dadurch wird die Beschreibung einer Valenzwellenfunktion mit Hilfe von Basisfunktionen genau so aufwendig wie die Darstellung der Wellenfunktion eines Core-Elektrons. Im Pseudoatom hingegen werden die oszillierenden Valenzwellenfunktionen durch glatte, nicht oszillierende Pseudowellenfunktionen ersetzt, die sich numerisch wesentlich leichter darstellen lassen. Ohne die Einführung dieser Pseudowellenfunktionen wären Untersuchungen von Festkörpern mit schweren Elementen mittels ebener Wellen oder mittels schwach lokalisierter Basisfunktionen nahezu undurchführbar.

Ein wesentliches Problem der Pseudopotentialkonstruktion besteht in der Frage, welche Elektronen als Valenzelektronen in das Pseudoatom übernommen werden müssen, und welche Elektronen als starre Core-Elektronen eliminiert werden können. Als erstes stellen wir fest, daß solch eine Zerlegung nicht in exakter Weise vorgenommen werden kann, sondern grundsätzlich zu einem systematischen Fehler führt, da insbesondere der Korrelationsanteil des Selbstenergie-Operators nicht linear mit den einzelnen Zuständen zusammenhängt. In der LDA besteht dieser Fehler darin, daß das Austausch-Korrelations-Potential additiv in zwei Funktionen zerlegt wird, obwohl es nicht linear in der Ladungsdichte ist:

$$\rho(\mathbf{r}) \quad = \quad \rho_{\mathrm{Core}}(\mathbf{r}) + \rho_{\mathrm{Val}}(\mathbf{r}) \quad ,$$

$$\text{aber} \quad V_{\text{xc}}(\rho(\mathbf{r})) \quad \neq \quad V_{\text{xc}}(\rho_{\text{Core}}(\mathbf{r})) + V_{\text{xc}}(\rho_{\text{Val}}(\mathbf{r})) \quad .$$

Der Fehler wird besonders groß, wenn ein starker räumlicher Überlapp zwischen der Coreladungsdichte $\rho_{\text{Core}}(\mathbf{r})$ und der Valenzladungsdichte $\rho_{\text{Val}}(\mathbf{r})$ besteht. Bei leichten Elementen aus der zweiten und dritten Periode des Periodensystems ist dieser Überlapp sehr gering, und die Aufteilung in Core- und Valenzladungsdichten ist ohne weiteres gerechfertigt. Bei schwereren Elementen nimmt der Überlapp zu, und es sind systematische Fehler in der Beschreibung von Festkörpern zu erwarten. Das trifft auch auf die leichteren Elemente der 1. Hauptgruppe (Alkalimetalle) zu. Wir merken an, daß dieses Problem durch Berücksichtigung nichtlinearer Core-Korrekturen (NLCC) teilweise überwunden werden kann [Lou82]. Das betrifft insbesondere die strukturellen Eigenschaften von Festkörpern [Eng90, Pal94, Zak95]. Für die Konstruktion des Selbstenergie-Operators und die Bestimmung von Bandstrukturen trifft das allerdings nur eingeschränkt zu.

Eine wesentliche Forderung an ein Pseudopotential ist die Übertragbarkeit auf eine geänderte chemische Umgebung des Atoms, wie sie im Kristall durch die Bindungspartner gegeben ist. Hier haben sich nichtlokale, normerhaltende ab-initio Pseudopotentiale [Ham79, Bac82, Kle82, Ham91] als besonders erfolgreich erwiesen. Das Konzept solcher Potentiale werden wir hier direkt auf die Beschreibung von Semicore- und Core-Elektronen anwenden. Wir folgen dabei den Konstruktionsvorschriften von Hamann und verwenden das von Hamann entwickelte Computerprogramm [Ham91]. Wir erläutern die Vorgehensweise am Beispiel des Cd Atoms. Seine Elektronenkonfiguration lautet $1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2$, oder kürzer $(\text{Ar})4s^2 4p^6 4d^{10} 5s^2$. In Abb. 12.1 stellen wir die LDA-Wellenfunktionen des Cd Atoms aus einer all-electron-Rechnung dar.

Wir konstruieren nun ein Pseudopotential (PP), das die atomaren Zustände $4s$, $4p$, $4d$ und $5s$ als Valenzelektronen beschreibt. Da dieses Pseudopotential 20 Valenzelektronen umfassen wird, bezeichnen wir es als ein Cd^{20+} PP. Bei herkömmlichen Pseudopotentialen stammen alle Valenzelektronen aus der gleichen Schale n. Jedem Valenzzustand (Drehimpuls l) wird eine Pseudowellenfunktion zugeordnet, deren Radialanteil knotenlos verläuft, und die dem untersten möglichen Zustand zu diesem Drehimpuls entspricht. Bei dem Cd^{20+} PP treten hingegen zwei s-Funktionen auf, nämlich der $4s$ Zustand und der $5s$ Zustand. Die entsprechenden Pseudowellenfunktionen können nicht beide knotenfrei verlaufen, denn sie müssen orthogonal zueinander sein. Die Pseudowellenfunktion zu $5s$ muß daher einen Knoten aufwei-

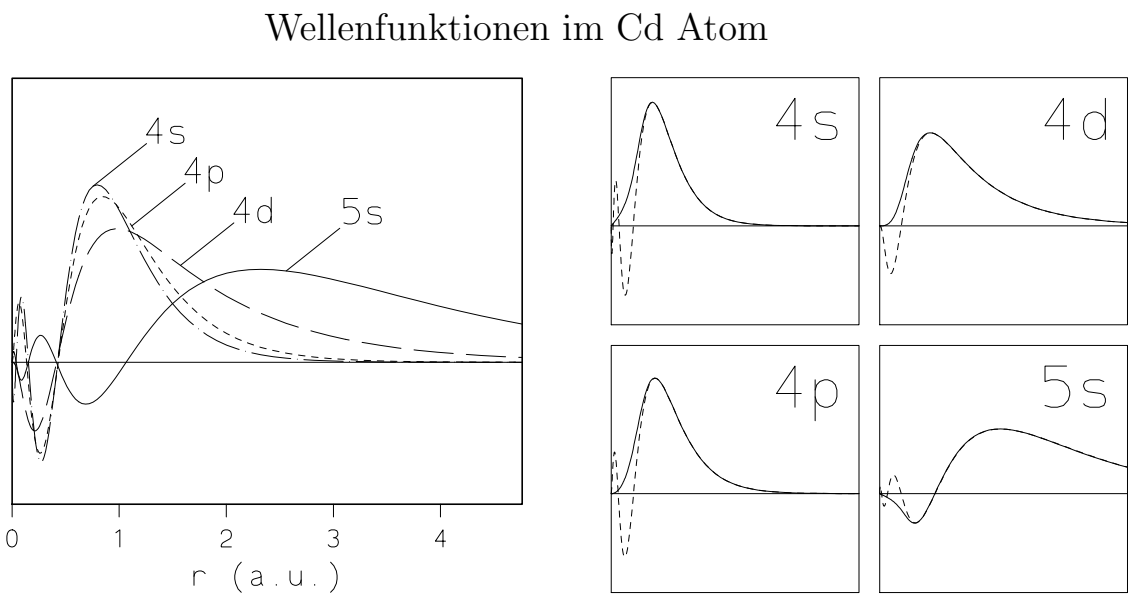


Abbildung 12.1: *Links: Radiale Wellenfunktionen $r\psi$ im Cd Atom, berechnet im Rahmen einer LDA all-electron Rechnung. Rechts: Vergleich der Pseudowellenfunktionen (durchgezogen) mit diesen all-electron Wellenfunktionen (gestrichelt). Es wurde das Cd^{20+} PP benutzt.*

sen. Es wäre kompliziert, unter der Randbedingung der Orthogonalität für beide Zustände gleichzeitig Pseudowellenfunktionen zu konstruieren. Wir schlagen statt dessen einen anderen Weg ein [KrüPr]. Aus dem neutralen Cd Atom werden zunächst die beiden $5s$ Elektronen entfernt. Für das verbleibende positiv geladene Cd^{2+} *Ion* mit der Konfiguration $(Ar)4s^2 4p^6 4d^{10}$ führen wir eine gewöhnliche Pseudopotential-Konstruktion durch. Dabei werden den Zuständen $4s$, $4p$ und $4d$ Pseudowellenfunktionen zugeordnet, und es werden Pseudopotentialanteile zu den Drehimpulsen l=0, 1 und 2 bestimmt. Dieses Pseudopotential wenden wir nun testweise auf das neutrale Cd *Atom* an. Dabei ergeben sich die Pseudozustände $4s$, $4p$ und $4d$ als erste Wellenfunktion zum jeweiligen Drehimpuls. Zusätzlich tritt der Zustand $5s$ auf; die zugehörige Wellenfunktion ist nun automatisch orthogonal zur $4s$ Pseudowellenfunktion.

Ein Pseudopotential muß in der Lage sein, bei der Anwendung auf das freie Atom die gleichen Energieeigenwerte und außerhalb des Cores die gleichen Wellenfunktionen zu liefern wie die all-electron-Rechnung. Im vorliegenden Fall ist das nicht ohne weiteres gewährleistet, denn das Pseudopotential wurde am geladenen Ion konstruiert und dann auf das neutrale Atom übertragen. Erfreulicherweise sind

Energien im Cd Atom

[eV]	all-electron	Pseudo-Atom
$4s$	–105.83	–105.73
$4p$	–66.99	–66.89
$4d$	–11.90	–11.84
$5s$	–5.96	–5.99

Tabelle 12.1: *LDA Energie-Eigenwerte der vier Zustände 4s, 4p, 4d und 5s des Cd Atoms. Erste Spalte: im Rahmen einer all-electron Rechnung bestimmt. Zweite Spalte: Mit Hilfe des Cd^{20+} PP berechnet.*

die Reproduktionsforderungen am neutralen Atom tatsächlich hervorragend erfüllt. In Abb. 12.1 vergleichen wir auf der rechten Seite die vier Pseudowellenfunktionen des Pseudoatoms mit den vier Wellenfunktionen aus der all-electron Rechnung. In allen Fällen stellen wir außerhalb des Corebereichs eine ausgezeichnete Übereinstimmung fest. Das gilt insbesondere auch für die $5s$ Wellenfunktion, die sich im Pseudoatom ohne weitere Randbedingung automatisch ergibt. Bei den LDA-Einteilchenenergien bietet sich eine ähnlich gute Übereinstimmung (siehe Tabelle 12.1). Die Abweichungen zwischen den Energien der all-electron Rechnung und denen der Pseudopotentialrechnung betragen maximal 0.1 eV. Sie kommen nicht allein durch die PP-Konstruktion selbst zustande, sondern auch durch die Fitprozedur, mit der die Pseudopotentiale in die von Bachelet et al. [Bac82] vorgeschlagene Form gebracht werden:

$$V_{\mathrm{PS}}^{(l)}(r) = -\frac{Z}{r}\left[c_1 \mathrm{erf}(\sqrt{\alpha_1} r) + c_2 \mathrm{erf}(\sqrt{\alpha_2} r)\right] + \sum_{i=1}^{N} \left(A_i^{(l)} + r^2 B_i^{(l)}\right) e^{-\gamma_i^{(l)} r^2} \quad . \quad (12.9)$$

Die in diesem Abschnitt vorgestellte Konstruktionsvorschrift läßt sich auf jedes gewünschte Pseudopotential anwenden. Zu jedem Drehimpuls berücksichtige man den jeweils tiefsten als Valenzzustand gewünschten Zustand. Alle höheren Zustände entferne man zunächst und konstruiere für das geladene Ion ein Pseudopotential, das man dann am neutralen Atom teste, bevor es auf den Festkörper übertragen wird.

Kapitel 13

Der Einfluß von Semicore-Elektronen auf die Bandstruktur

Wir wenden uns nun der Fragestellung zu, in welcher Form die Semicore-Elektronen die Festkörpereigenschaften eines Kristalls, insbesondere die Bandstruktur, beeinflussen. Sollte sich dieser Einfluß als wesentlich herausstellen, so können die Semicore-Zustände nicht als Core-Zustände im Pseudopotential (PP) Zugang eliminiert werden. Die PP-Methode beruht auf den folgenden Annahmen für die eliminierten Zustände:
(i) Die Eigenschaften der eliminierten Zustände ändern sich nicht, wenn das Atom in eine neue chemische Umgebung gebracht wird.
(ii) Die Auswirkungen der eliminierten Zustände auf die Eigenschaften des Festkörpers können über das Pseudopotential exakt beschrieben werden.
Die zweite Annahme muß sofort in Frage gestellt werden, wenn auf der Basis von LDA-Pseudopotentialen eine GW-Rechnung durchgeführt werden soll. Im Rahmen der LDA wird vom Austausch-Korrelations (XC) Potential aller Elektronen im Atom das XC-Potential der atomaren Valenzelektronen subtrahiert (Pseudopotential-Konstruktion) und anschließend das XC-Potential der Festkörper-Valenzelektronen addiert:

$$V_{\mathrm{xc}} = \underbrace{\left(V_{\mathrm{xc}}^{\mathrm{Atom,all}} - V_{\mathrm{xc}}^{\mathrm{Atom,Val}}\right)}_{\mathrm{(PP)}} + V_{\mathrm{xc}}^{\mathrm{Fk,Val}} \quad . \tag{13.1}$$

Diese Aufteilung ist gerechtfertigt, wenn die Valenz-Ladungsdichteverteilung im Bereich der Core-Elektronen sich im Atom und im Festkörper nicht wesentlich unter-

scheiden. Wenn nun auf der Basis solcher LDA-Pseudopotentiale eine GW-Rechnung für den Festkörper durchgeführt wird, dann entspricht das dem folgenden Ausdruck für den gesamten Selbstenergie-Operator:

$$\Sigma_{\mathrm{xc}} \sim \underbrace{\left(V_{\mathrm{xc}}^{\mathrm{Atom,all}} - V_{\mathrm{xc}}^{\mathrm{Atom,Val}}\right)}_{\mathrm{(PP)}} + \Sigma_{\mathrm{xc}}^{\mathrm{Fk,Val}} \quad . \tag{13.2}$$

Dabei tritt aber ein systematischer Fehler auf, denn der richtige Ausdruck für eine Aufteilung des Selbstenergie-Operators müßte lauten:

$$\Sigma_{\mathrm{xc}} = \left(\Sigma_{\mathrm{xc}}^{\mathrm{Atom,all}} - \Sigma_{\mathrm{xc}}^{\mathrm{Atom,Val}}\right) + \Sigma_{\mathrm{xc}}^{\mathrm{Fk,Val}} \quad . \tag{13.3}$$

Zwischen den beiden Klammerausdrücken in Gl. (13.2) und (13.3) besteht ein fundamentaler Unterschied:

$$\Delta\Sigma_{\mathrm{xc}} = \left(\Sigma_{\mathrm{xc}}^{\mathrm{Atom,all}} - \Sigma_{\mathrm{xc}}^{\mathrm{Atom,Val}}\right) - \left(V_{\mathrm{xc}}^{\mathrm{Atom,all}} - V_{\mathrm{xc}}^{\mathrm{Atom,Val}}\right) \neq 0 \quad . \tag{13.4}$$

Die Differenz $\Delta\Sigma_{\mathrm{xc}}$ kommt dadurch zustande, daß *beim Übergang von der LDA zur GWA die Methode zur Beschreibung von Austausch und Korrelation gewechselt wird, das LDA Pseudopotential aber beibehalten wird.* Die Differenz geht direkt in die elektronische Quasiteilchen-Bandstruktur ein und verfälscht sie. Selbst wenn die eingangs geforderten zwei Bedingungen (i) und (ii) im Rahmen der LDA erfüllt sind, so bedeutet das noch lange nicht, daß auch die Differenz $\Delta\Sigma_{\mathrm{xc}}$ verschwindet. Sie kann allerdings systematisch reduziert werden, indem sukzessive die Zahl der explizit berücksichtigten Valenzelektronen vergrößert wird. Dabei werden beide Klammerausdrücke in Gl. (13.4) verkleinert (im Grenzfall einer all-electron Rechnung zu Null), und letztlich auch ihre Differenz. Diese systematische Vergrößerung der Zahl explizit berücksichtigter Valenzelektronen werden wir im folgenden untersuchen. Das bedeutet letztlich einen gezielten Schritt auf dem Weg zu einer all-electron Rechnung. Wir werden diesen Weg aber nur so weit beschreiten, wie es zur konvergenten Behandlung der zu untersuchenden Fragestellung notwendig ist. Dadurch wird der Rechenaufwand auf das notwendige Maß beschränkt. Die größten Änderungen in der Valenzbandstruktur sind dann zu erwarten, wenn Elektronen der obersten vollständig besetzten Schale in die Valenzschale überführt werden. Diese Elektronen, die am ehesten einen Einfluß auf die „konventionellen“ Valenzelektronen auszuüben vermögen, bezeichnen wir als *Semicore-Elektronen.* Die Effekte durch Einbeziehen noch tieferer Core-Elektronen werden eher gering ausfallen. Um die grundlegenden Eigenschaften dieser Einflüsse zu studieren, beschränken wir uns in diesem Kapitel auf die systematische Einbeziehung der Semicore-Elektronen allein.

13.1 CdS als Prototyp

II-VI Halbleiter zählen zu den Materialien, bei denen die Effekte der Semicore-Elektronen am deutlichsten zu Tage treten. Die Energien der Semicore d-Elektronen der Kationen (Zn, Cd) liegen im Bereich der Valenzbänder, und die d-Elektronen wechselwirken mit den Valenzelektronen. Es ist seit langem bekannt, daß diese d Elektronen in Form von Valenzelektronen Berücksichtigung finden müssen, um die strukturellen Eigenschaften der Materialien im Rahmen der LDA konsistent beschreiben zu können [Wan81, Mar91, Hua93, Sch93a, Sch93b, Xu93]. Das geschieht mit Hilfe kationischer (12+)-Pseudopotentiale, die neben den s-Valenzelektronen auch die d-Zustände der Semicore-Schale als Valenzelektronen behandeln.

Für eine Untersuchung des Einflusses solcher Elektronen auf die strukturellen und elektronischen Eigenschaften von halbleitenden Kristallen bieten sich die II-VI Halbleiter geradezu an. In keiner anderen Materialklasse besitzen diese Zustände eine so hohe Bandstruktur-Energie in den Valenzbändern [abgesehen von den Übergangselementen, die aber ganz andere physikalische Eigenschaften aufweisen als die IV-, III-V und II-VI Halbleiter.] Demzufolge ist ihr Einfluß auf die Festkörpereigenschaften bei den II-VI Verbindungen besonders groß. Außerdem sind die d Zustände in den Elementen der Gruppen III, IV und höheren immer stärker lokalisiert, was ihre numerische Beschreibung erheblich erschwert. Beide Eigenschaften zusammen – die hohe Energie und die moderate Lokalisierung – stoßen gewissermaßen bei den II-VI Verbindungen das Tor zur Physik der kationischen d Zustände auf.

Die Fragestellung nach dem Einfluß der d-Zustände auf die Eigenschaften der II-VI Halbleiter wurde mit Hilfe von kationischen 2+ und 12+ PP'en im Rahmen der LDA untersucht [Sch93a]. Dabei wurden (12+)-PP LDA-Bandstrukturen bestimmt. An diesen LDA-Bandstrukturen setzt unsere Untersuchung an. Die Bandstrukturen weisen zwei wesentliche Unzulänglichkeiten auf:
(i) Das fundamentale Gap ist im Vergleich zum Experiment viel zu klein.
(ii) Die energetische Position der Semicore d Elektronen ist im Vergleich zum Experiment zu hoch.
In der Unterschätzung des Gaps erkennen wir das generelle Gap-Problem der LDA wieder. Aufgrund der Erfahrungen mit anderen Materialien könnte man erwarten, daß Quasiteilchen-Korrekturen im Rahmen der GWA auch hier zu einer richtigen Beschreibung der Energielücke führen werden. Auch das zweite Problem ist uns nicht neu. Wir haben bei der Untersuchung von Oberflächenzuständen an H-terminierten

Halbleiter-Oberflächen gesehen, daß die LDA für Zustände mit stark lokalisierten Wellenfunktionen zu hohe Bandstruktur-Energien ergibt. Durch Quasiteilchen-Korrekturen hatten wir eine deutliche Energieabsenkung erzielt und eine hervorragende Übereinstimmung mit dem Experiment erreicht. Die kationischen d-Elektronen der II-VI Halbleiter sind ähnlich stark lokalisiert wie Wasserstoff-Zustände. Man könnte daher auch für sie Energieabsenkungen im eV-Bereich erwarten. Bei der Durchführung einer Standard-GWA Rechnung für CdS (unter Einbeziehung der kationischen d-Elektronen) stellen wir allerdings fest, daß das Gap nicht wesentlich vergrößert wird, und daß die d-Elektronen eine erhebliche Energie-*Anhebung* anstelle einer Absenkung erfahren.

Um diese Effekte verstehen zu können, und um letztlich zu einer Überwindung dieser Artefakte zu gelangen, stellen wir hier am Beispiel CdS eine systematische Studie unter Verwendung verschiedener Pseudopotentiale (von Cd^{2+} bis Cd^{20+}) mit einer wachsenden Zahl von Valenzelektronen vor. Wir beschränken uns dabei auf die kubische Modifikation (Zinkblende-Struktur), die numerisch weniger aufwendig zu handhaben ist als die hexagonale Wurtzit-Modifikation. Die grundlegenden Erkenntnisse, die wir daraus gewinnen, gelten gleichermaßen für beide Modifikationen. Bei den Rechnungen erweist sich die Verwendung von Gauß-Funktionen als Basissatz sowohl für die LDA-Wellenfunktionen als auch für die GWA-Zweipunktfunktionen als unschätzbarer Vorteil. Wir benötigen lediglich 90 Ansatzfunktionen für die LDA bzw. 190 für den GW-Basissatz. Bei Verwendung ebener Wellen müßte man mehrere tausend Basisfunktionen benutzen, und die Rechnungen wären um mindestens eine Größenordnung aufwendiger. Die Ladungsdichte und das LDA-Potential stellen wir in diesem Abschnitt allerdings noch nicht mit Gauß-Orbitalen dar (wie in Abschnitt 12.1 beschrieben), sondern nach wie vor mit ebenen Wellen. Dadurch erreichen wir auch bezüglich der totalen Energie die zur Bestimmung der Gitterkonstante notwendige Genauigkeit. In der GWA berechnen wir die statische dielektrische Funktion im Rahmen der RPA und erweitern sie mittels des gewöhnlichen Plasmon-Pol Modells (I) zur dynamischen Matrix. Wir verwenden keine dielektrischen Modellfunktionen, denn es ist nicht klar, ob sie auch bei Berücksichtigung der stark lokalisierten Semicore-Zustände eine richtige Beschreibung der Abschirmeffekte liefern.

Wir beginnen die Untersuchung mit einer Rechnung, in der nur die Cd $5s$ Elektronen als Valenzelektronen eingehen. Danach werden sukzessive die $4d$, $4p$ und $4s$ Elektronen als Valenzelektronen mit berücksichtigt. Wir fassen in Tabelle 13.1 die entsprechenden Pseudopotentiale zusammen. Als Cd^{2+} PP und Cd^{12+} PP verwenden

Cd PP	Valenzelektronen			
	$4s$	$4p$	$4d$	$5s$
2+				×
12+			×	×
18+		×	×	×
20+	×	×	×	×

Tabelle 13.1: *Struktur der vier Cd Pseudopotentiale.*

wir die von Bachelet et al. konstruierten Potentiale [Bac82]. Die Potentiale Cd^{18+} und Cd^{20+} konstruieren wir selber nach der Methode von Hamann [Ham91] (siehe Abschnitt 12.2). In allen Rechnungen beschreiben wir das Schwefel-Atom mit dem S^{6+} PP von Bachelet et al. [Bac82]. Die Semicore-Elektronen des Schwefels haben wegen ihrer weit stärkeren Lokalisierung nur geringen Einfluß auf die Bandstruktur und werden in unseren Rechnungen nicht als Valenzelektronen berücksichtigt.

Wenden wir uns zunächst den Rechnungen mit dem Cd^{2+} PP zu. Wir erhalten keine gute Übereinstimmung der berechneten mit experimentellen Daten. Die theoretische Gitterkonstante, die wir im Rahmen der LDA berechnen, ist mit 5.05 Å viel kleiner als die experimentelle Gitterkonstante (5.82 Å). Die LDA- und GW-Bandstrukturen dieser Rechnung stellen wir in Abb. 13.1 (links) dar. [Wir berechnen alle Bandstrukturen in diesem Kapitel bei der experimentellen Gitterkonstante, um die Ergebnisse der verschiedenen Potentiale miteinander vergleichen zu können.] In der LDA erhalten wir ein fundamentales Energiegap von 2.16 eV. Es wird durch die Quasiteilchen-Korrekturen auf 3.70 eV angehoben. Dieses Gap ist weit größer als die experimentellen Werte. Cardona et al. haben kubische CdS-Filmen auf GaAs aufgewachsen und daran mittels Reflektivitäts-Messungen bzw. optischer Ellipsometrie ein fundamentales Gap von 2.50 eV bzw. 2.55 eV ermittelt [Car65]. Neuere Ellipsometrie-Messungen von Rossow et al. und von Zahn et al. an kubischen CdS-Filmen auf InP beziffern das Gap auf 2.41 eV [Ros93] bei RT bzw. auf 2.48 eV [Zah94] bei einer Temperatur von 1.6 K.

Die Unzulänglichkeiten in diesem Cd^{2+} PP Zugang überraschen nicht. Ihre Ursache besteht in der inadäquaten Behandlung der Cd $4d$ Zustände als Core-Zustände. Ihre LDA-Energie beträgt im Atom –11.90 eV und liegt somit zwischen den atomaren Schwefel-Niveaus $3s$ (–17.28 eV) und $3p$ (–7.11 eV). In CdS kommt es zu einer erheb-

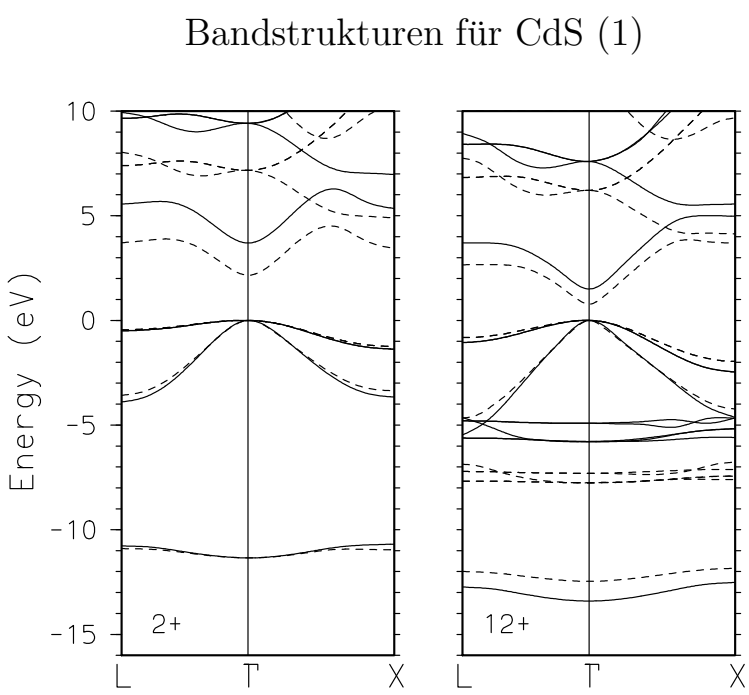


Abbildung 13.1: *Berechnete Bandstrukturen für kubisches CdS. Durchgezogene Kurven: GWA Quasiteilchen-Bandstruktur. Gestrichelt: LDA-Energien. Es werden die Ergebnisse dargestellt, die mit dem Cd^{2+} PP (links) sowie mit dem Cd^{12+} PP (rechts) berechnet wurden.*

lichen Wechselwirkung zwischen den Valenz-Zuständen (die hauptsächlich von den Schwefel-Orbitalen gebildet werden) und den Cd 4*d* Zuständen. Die Cd 4*d* Wellenfunktionen werden dadurch in ihrer räumlichen Struktur erheblich verändert (siehe Abb. 13.2), und eine Behandlung der Zustände als starre Core-Zustände im Cd^{2+} PP erweist sich als unzureichend. Es kommt insbesondere zu einer *p*-*d* Kopplung zwischen den S 3*p* und den Cd 4*d* Zuständen.

Wir beziehen im nächsten Schritt die Cd 4*d* Elektronen in die Rechnungen mit ein. Das geschieht durch Verwendung des Cd^{12+} PP. Bei der Untersuchung der Gitterkonstante im Rahmen der LDA stellen wir eine bedeutende Verbesserung fest. Die Gitterkonstante beträgt nunmehr 5.61 Å und ist nur noch 4 % kleiner als der experimentelle Wert. Das LDA-Gap erfährt eine deutliche Absenkung und beträgt nun 0.78 eV (siehe die Bandstruktur in Abb. 13.1, rechte Seite) Wir merken an, daß die Berücksichtigung nichtlinearer Core-Korrekturen (NLCC; vgl. Abschnitt 12.2) eine ähnliche Verkleinerung des LDA-Gaps bewirkt [Zak94].

Das Quasiteilchen-Gap beträgt in unserer GW-Rechnung 1.50 eV. Dieser Wert

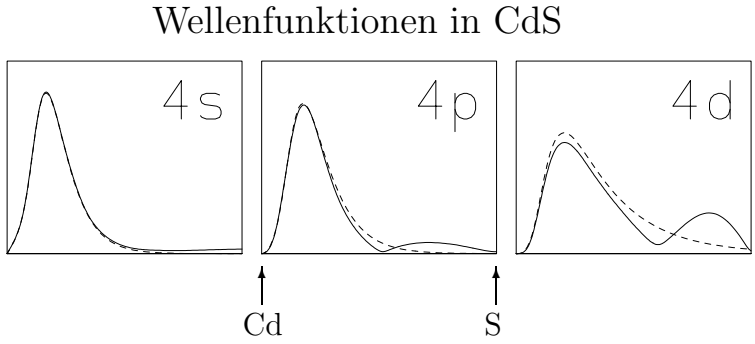


Abbildung 13.2: *Beträge der Wellenfunktionen der Cd Semicore-Elektronen in kubischem CdS (durchgezogene Kurven) im Vergleich zu den entsprechenden atomaren Wellenfunktionen im Cd Pseudo-Atom (gestrichelt). Die Darstellung der Festkörper-Wellenfunktionen erfolgt auf einer Cd-S-Bindung entlang der [111]-Richtung. Am linken Kastenrand befindet sich jeweils das Cd Atom, am rechten das S-Atom (durch Pfeile angedeutet).*

liegt zwar um 0.8 eV über dem LDA-Ergebnis, aber um 1 eV unter dem experimentellen Wert. Für ein GW-Quasiteilchen-Ergebnis ist das eine ungewöhnlich große Abweichung. Die erstmalig auftretenden Cd 4*d* Elektronen haben eine mittlere Bandstrukturenergie von –7.4 eV in LDA. Im Experiment beobachtet man hingegen weit niedrigere Energien von –9.5 eV [Ley74, Sto83]. Winkelaufgelöste Photoemissions-Spektroskopie (ARPES) Untersuchungen von Stampfl et al. ergeben Energien zwischen –9.5 und –8.9 eV an verschiedenen Punkten der Brillouinzone [Sta96], mit einem Mittelwert von –9.2 eV.

An dieser Stelle erscheint uns eine grundsätzliche Anmerkung zur Aufspaltung der Cd 4*d* Zustände wichtig, um eventuellen Mißverständnissen vorzubeugen. In einer LDA-Rechnung zum Cd Atom ohne Berücksichtigung der Spin-Bahn-Wechselwirkung sind die fünf Energieniveaus der 4*d* Zustände energetisch entartet, weil das Atom Kugelsymmetrie aufweist. Durch die andere chemische Umgebung der Cd Atome in CdS wird die Kugelsymmetrie durchbrochen, und es kommt zu einer Aufhebung der energetischen Entartung der Zustände. Ihre Aufspaltung liegt in der Größenordnung 0.5 eV. Diese Aufspaltung darf nicht mit der Spin-Bahn Aufspaltung der Zustände verwechselt werden, die im Atom 0.7 eV beträgt [Lan50] und somit in der gleichen Größenordnung liegt. Sie wird in unseren Rechnungen nicht explizit berücksichtigt. Durch die Spin-Bahn Wechselwirkung kommt es zusätzlich zu der geometriebedingten Kristallfeld-Aufspaltung noch zu einer weiteren Aufspaltung der Niveaus, die in unseren Bandstrukturen nicht enthalten ist. Das gilt auch

für alle weiteren Bandstrukturen und lokalisierten Zustände, die wir in dieser Arbeit behandeln. Falls wir die Spin-Bahn Wechselwirkung explizit berücksichtigen, weisen wir an der entsprechenden Stelle ausdrücklich darauf hin.

Wenden wir uns nun den Quasiteilchen-Effekten der Cd $4d$ Zustände in der Cd^{12+} PP Rechnung zu. Vor dem Hintergrund der Erfahrungen mit lokalisierten Wasserstoff-Zuständen an H-terminierten Oberflächen könnte man erwarten, daß die Cd $4d$ Energieniveaus durch GW Quasiteilchen-Korrekturen zu tieferen Energien abgesenkt werden. *Wir beobachten im Gegenteil, daß die Cd 4d Elektronen um mehr als 2 eV zu höheren Energien verschoben werden.* Dieses Ergebnis der GW-Rechnung für das Cd^{12+} PP steht in klarem Widerspruch zum Experiment. Die Quasiteilchen-Energien der Cd $4d$ Elektronen liegen um 4 eV über den gemessenen Werten. Dieses Ergebnis wurde von Fleszar et al. in ihrer GW-Rechnung bestätigt [Fle96]. Bisher wurde für kein Material eine derart große Quasiteilchen-Korrektur eines besetzten Zustands zu höheren Energien beobachtet. Weiterhin stellen wir fest, daß die Störmatrix $(\Sigma - V_{\mathrm{xc}})_{mm'}$ (vgl. Gl. (2.13)) in unserer Rechnung nicht diagonal ist. Es treten Nichtdiagonalelemente zwischen den Cd $4d$ Zuständen und den S $3p$-artigen höheren Valenzzuständen auf. Daraus ergeben sich Störbeiträge in zweiter Ordnung zu den Energien und Beiträge in erster Ordnung zu den Wellenfunktionen. Für die Cd $4d$ und die S $3p$ Zustände stimmen also die Wellenfunktionen ψ^{LDA} und ψ^{GWA} nicht überein, und die p-d-Wechselwirkung wird im Rahmen der GWA nicht richtig beschrieben. Zakharov et al. stellten an ZnS den gleichen Effekt fest [Zak94]. Auch dieses Resultat ist sehr ungewöhnlich, denn in den allermeisten GW-Rechnungen stellt sich der Störoperator als nahezu diagonal heraus, und die LDA- und GWA-Wellenfunktionen sind gewöhnlich zu mehr als 99 % identisch [Hyb86, God88].

Die unzureichende Beschreibung der Bandstruktur von CdS ist eine direkte Folge der Verwendung des Cd^{12+} PP. Die Aufteilung der Cd Semicore-Schale in $4s$ und $4p$ Core-Elektronen einerseits und $4d$ Valenzelektronen andererseits erweist sich als unzulässig. Im Rahmen der LDA erscheint diese Aufteilung wegen der großen Bindungsenergie von 106 eV bzw. 67 eV der $4s$ und $4p$ Elektronen gerechtfertigt, die eine Wechselwirkung mit den Valenzelektronen weitgehend ausschließt. Dabei wird jedoch nicht beachtet, daß die $4s$ und $4p$ Elektronen in einem ähnlichen Raumbereich um den Cd Kern lokalisiert sind wie die $4d$ Elektronen (vgl. Abb. 12.1). Für alle drei Zustände nehmen die Wellenfunktionen bei einem Radius von etwa 1 a.u. ihr Maximum an. Daraus resultiert ein starker räumlicher Überlapp der Wellen-

funktionen und eine erhebliche Wechselwirkung zwischen den Elektronen der Schale. Das gilt besonders für die Austausch-Wechselwirkung, in die die Energien der beteiligten Zustände gar nicht eingehen, sondern nur ihre räumliche Struktur. Wenn die Semicore-Schale in Core- und Valenzelektronen unterteilt wird, so ist beim Übergang von der LDA zur GWA eine große, unkontrollierbare Differenz zwischen den nichtlinearen Austausch-Korrelations-Effekten der beiden Methoden zu erwarten. Die in Gl. (13.4) dargestellte Differenz verschwindet nicht, sondern führt zu den starken Abweichungen der GW-Ergebnisse von bis zu 4 eV von den experimentellen Daten. Wir erwarten derartige Probleme auch bei anderen Materialien, wenn Zustände einer Schale trotz ähnlicher Lokalisierungs-Eigenschaften in Core- und Valenz-Zustände unterteilt werden.

Man gelangt zu einer verläßlicheren Beschreibung von CdS im Rahmen der GWA, wenn die $4p$ und $4s$ Zustände als Valenzelektronen mitgenommen werden, denn die Differenz (13.4) wird dabei reduziert. Wir führen dies in zwei Schritten durch. Zunächst beziehen wir die Cd $4p$ Zustände in das Pseudopotential mit ein und konstruieren ein Cd^{18+} PP. Die $4s$ Elektronen verbleiben nach wie vor im Core. Aus Abb. 13.2 geht hervor, daß die $4p$ Elektronen durch den Kristall nur eine geringe Verformung erfahren. Dementsprechend gering sind die damit verbundenen Änderungen der LDA-Ergebnisse gegenüber der Cd^{12+} Rechnung, in der die $4p$ Zustände als Core-Zustände per Konstruktion die gleichen waren wie im Cd Atom. Durch die jetzt vorhandene Flexibilität der $4p$ Elektronen ergibt sich eine etwas größere Gitterkonstante von 5.72 Å. In der LDA-Bandstruktur (siehe Abb. 13.3, links) betragen die Veränderungen weniger als 0.1 eV für die Valenzzustände und weniger als 0.3 eV für Leitungszustände. In der GW Quasiteilchen-Bandstruktur bietet sich hingegen ein ganz anderes Bild. Die unvernünftig große Verschiebung der Cd $4d$ Elektronen zu höheren Energien, die in der Cd^{12+} PP Rechnung auftrat, findet in dieser Cd^{18+} PP Untersuchung nicht mehr statt. Außerdem ist der Störoperator $\Sigma - V_{\mathrm{xc}}$ nun diagonal, und die LDA-Wellenfunktionen stimmen mit den GWA-Wellenfunktionen überein. Die p-d Wechselwirkung, die in der Cd^{12+} PP Rechnung innerhalb der GWA falsch beschrieben wurde, wird nun in dieser Rechnung besser erfaßt. Die falschen Ergebnisse für die p-d Wechselwirkung, die offensichtlich falsche Quasiteilchen-Korrektur der Cd $4d$ Elektronen und die Unterschiede zwischen LDA- und GWA-Wellenfunktionen im Rahmen der Cd^{12+} PP Rechnung erweisen sich als Artefakt der Pseudopotential-Methode. Wir stellen allerdings fest, daß das fundamentale Gap der Quasiteilchen-Bandstruktur mit 1.60 eV auch in der Cd^{18+} PP

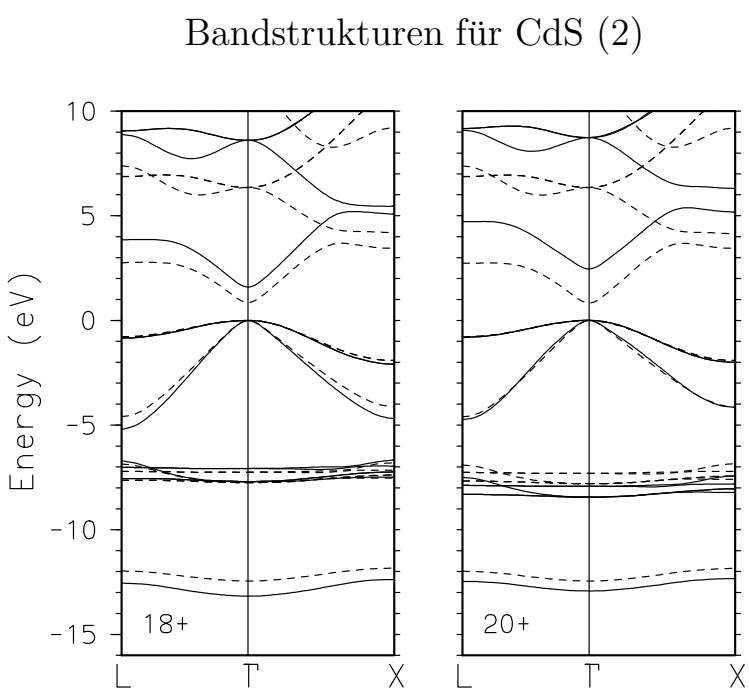


Abbildung 13.3: *Bandstrukturen für kubisches CdS (vgl. Abb. 13.1), berechnet mit dem Cd^{18+} PP (links) sowie mit dem Cd^{20+} PP (rechts). Durchgezogene Kurven: GWA Quasiteilchen-Bandstruktur. Gestrichelt: LDA-Energien.*

Rechnung noch deutlich unter dem experimentellen Wert liegt.

Wir schließen im letzten Schritt auch die Cd 4*s* Elektronen in die Valenzzustände mit ein und konstruieren ein Cd^{20+} PP. Die Gitterkonstante bleibt mit 5.73 Å gegenüber der Cd^{18+} PP Rechnung fast unverändert. Die verbleibende Abweichung um 1.5 % vom experimentellen Wert ist typisch für LDA-Rechnngen unter Verwendung der Austausch-Korrelations-Daten von Ceperley und Alder. In der LDA-Bandstruktur (siehe rechte Seite von Abb. 13.3) sind praktisch keine Abweichungen von der Cd^{18+} PP Rechnung festzustellen. Diese Ähnlichkeit der LDA-Ergebnisse mit denen des Cd^{18+} PP sind auch zu erwarten, denn der Cd 4*s* Zustand erfährt praktisch keine Veränderung gegenüber der atomaren Wellenfunktion (siehe Abb. 13.2). Wenn man sich auf LDA-Rechnungen beschränkt, so ist es ohne weiteres gerechtfertigt, diesen Zustand als Core-Zustand zu behandeln und per Pseudopotential-Konstruktion zu eliminieren. Hingegen ergeben sich in der GW Quasiteilchen-Bandstruktur nochmals deutliche Veränderungen gegenüber der Cd^{18+} PP Rechnung. Die Cd 4*d* Niveaus erfahren nun eine Quasiteilchen-Korrektur von –0.7 eV und haben damit eine mittlere Bandstruktur-Energie von –8.1 eV. Dieser Wert liegt immer noch um etwa

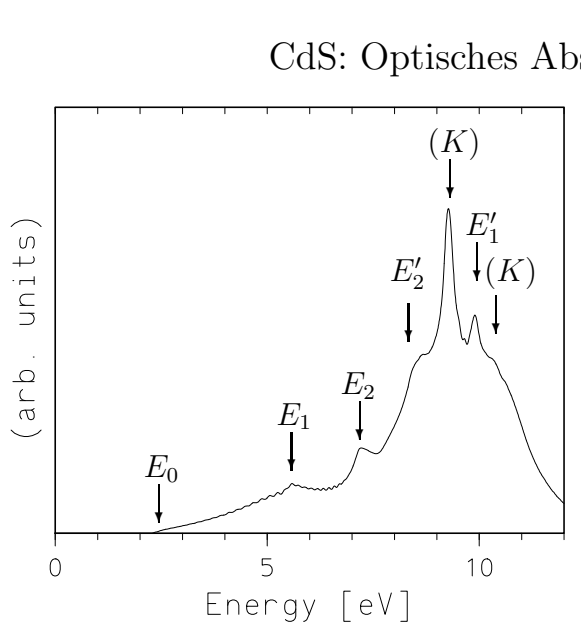
1 eV über dem experimentellen Resultat. Die verbleibende Abweichung hängt mit der grundlegenden Frage zusammen, inwieweit die Standard-GWA in der Lage ist, die Energie stark lokalisierter Zustände zu beschreiben. Wir kommen in Kap. 14 auf dieses Problem zurück. Der Störoperator $(\Sigma - V_{\rm xc})$ bleibt auch in dieser Rechnung diagonal. Das unterste Leitungsband wird gegenüber der Cd^{18+} PP GW-Rechnung nochmals deutlich um 0.85 eV angehoben. Das endgültige Quasiteilchen-Gap beträgt nunmehr 2.45 eV und stimmt mit dem gemessenen Gap von 2.48 eV hervorragend überein.

Ein Großteil der experimentellen Daten zur Bandstruktur von kubischem CdS wurden aus optischen Messungen, besonders aus Ellipsometrie- und Reflektivitäts-Untersuchungen, gewonnen [Car65, Hof93, Ros93, Zah94]. Um mit den experimentellen Spektren direkt vergleichen zu können, bestimmen wir das optische Absorptionsspektrum [Iba95]

$$A(\omega) = A_0 \cdot \frac{1}{\omega^2} \sum_{\rm m}^{\rm occ} \sum_{\rm n}^{\rm empty} \sum_{\mathbf{k}} |\langle {\rm m}\mathbf{k}|\hat{\mathbf{p}}|{\rm n}\mathbf{k}\rangle|^2 \, \delta[\omega - (E_{{\rm n}\mathbf{k}} - E_{{\rm m}\mathbf{k}})] \quad . \tag{13.5}$$

[$\hat{\mathbf{p}}$ ist der Impulsoperator.] Wir werten diesen Ausdruck mit 11000 $\mathbf{k}$-Punkten aus. An jedem $\mathbf{k}$-Punkt bestimmen wir zunächst die LDA-Energien. Um die zugehörigen GW Quasiteilchen-Energien zu erhalten, haben wir aus den Quasiteilchen-Korrekturen an einigen Hochsymmetriepunkten ein Bandstrukturenergie-abhängiges Modell für die Korrektur konstruiert, das sich für alle $\mathbf{k}$-Punkte verwenden läßt. Die optischen Übergangsmatrix-Elemente in (13.5) berechnen wir mit den LDA-Wellenfunktionen. Das resultierende Spektrum wird in Abb. 13.4 dargestellt. Maxima in solch einem Absorptionsspektrum sind hauptsächlich von solchen $\mathbf{k}$-Punkten zu erwarten, an denen besetzte und unbesetzte Bänder parallel verlaufen (van Hove-Singularitäten). Das ist in der Regel nur an den Hochsymmetriepunkten der Fall. Diese liefern die markantesten Beiträge zum Spektrum. Dabei ist es von entscheidender Bedeutung, wie groß das Übergangs-Matrixelement $\langle {\rm m}\mathbf{k}|\hat{\mathbf{p}}|{\rm n}\mathbf{k}\rangle$ ist. Wenn der Übergang aus Symmetriegründen verboten ist, so ist das Matrixelement Null, und der Beitrag zum Spektrum verschwindet.

Neben dem Einsetzen der optischen Absorption bei E_0=2.45 eV finden wir im Spektrum bis 12 eV vier Maxima sowie zwei Schultern. Der erste Peak bei 5.52 eV (E_1) stammt vom niedrigsten Übergang am L-Punkt. Er wird im Experiment bei Energien zwischen 5.0 eV und 5.49 eV beobachtet [Car65, Hof93, Ros93]. Ein weiterer Peak bei 7.18 eV wird vom niedrigsten Übergang am X-Punkt hervorgerufen



[eV]		GWA	Exp.
E_0	(Γ)	2.45	2.50[a], 2.55[a], 2.48[c]
E_1	(L)	5.52	5.49[a], 5.3[a], 5.0[b], 5.1[d]
E_2	(X)	7.18	7.95[a], 7.5[a], 7.4[b]
E_2'	(X)	8.32	8.3[b]
	(K)	9.26	9.18[a], 8.8[a]
E_1'	(L)	9.98	
	(K)	10.35	

[a]Cardona et al. [Car65]
[b]Hofmann et al. [Hof93]
[c]Zahn et al. [Zah94]
[c]Rossow et al. [Ros93]

Abbildung 13.4: *Optisches Spektrum von kubischem CdS auf der Basis der Quasiteilchen-Bandstruktur (siehe Text). Die Bezeichnungen entsprechen der üblichen Nomenklatur (siehe z.B. [Hof93]).*

(E_2). In diesem Energiebereich treten in den Messungen ebenfalls Übergänge auf [Car65, Hof93]. Allerdings gestaltet sich die Zuordnung der Strukturen zueinander in diesem Energiebereich zunehmend schwierig. Beispielsweise treten in den Messungen von Hofmann et al. *zwei* Maxima bei 6.6 eV und 7.4 eV auf, während wir in unserem berechneten Spektrum nur *einen* Peak bei 7.18 eV finden. Die Schulter bei 8.32 eV entspricht dem zweitniedrigsten Übergang am X-Punkt (E_2'). Bei einer Energie von 8.74 eV befindet sich der zweite Übergang E_0' am Γ-Punkt. Sein Übergangs-Matrixelement ist jedoch sehr klein. Der Übergang ist daher im Spektrum nicht wahrzunehmen. Hingegen tritt bei 9.26 eV ein deutlicher Peak auf, der einem Übergang am K-Punkt zuzuordnen ist. Cardona et al. haben einen Übergang bei einer Energie von 9.18 eV gefunden [Car65], die mit unserem Ergebnis hervorragend übereinstimmt. Die Autoren interpretieren diesen Peak als E_1' Übergang (zweiter Übergang am L-Punkt). Wir finden den E_1' Übergang in unserer Quasiteilchen-Bandstruktur jedoch erst bei 9.98 eV.

Die Unterschiede zwischen dem von uns berechneten Spektrum und den experimentellen Resultaten können vielerlei Ursachen haben. Auf experimenteller Seite besteht eine gewisse Problematik sicherlich in der Probenpräparation. Als freier Kristall läßt sich CdS nicht in der kubischen Modifikation herstellen. Kubisches CdS

CdS: Quasiteilchen-Energien

[eV]	GWA	Exp.
Γ_{15}	0.0	0.0
$4d$ (Γ)	–8.2	–9.2
X_5	–2.01	–1.54
X_3	–4.17	–3.7
$4d$ (X)	–7.9	–9.1
X_1	–12.34	–13.85
L_3	–0.81	–0.62
L_1	–4.75	–4.18
$4d$ (L)	–8.0	–9.5

Tabelle 13.2: *Berechnete GW Quasiteilchen-Energien für kubisches CdS (Cd^{20+} PP Rechnung) im Vergleich zu ARPES-Daten von Stampfl et al. [Sta96].*

entsteht bei epitaktischem Wachstum auf einer geeigneten Oberfläche eines anderen kubischen Materials. Hier haben sich besonders die As-terminierte GaAs($1\bar{1}1$) Oberfläche [Car65] und die InP(110) Oberfläche [Hof93, Ros93, Zah94] bewährt. Auf diese Weise lassen sich Schichten von kubischem CdS präparieren. Typische Schichtdicken betragen einige 100 nm. Die Differenzen zwischen der Gitterkonstante von kubischem CdS (5.82 Å) und der des jeweiligen Substrats (GaAs: 5.65 Å; InP: 5.87 Å) führen in der CdS Schicht zu Verzerrungen, die die Bandstruktur beeinflussen können. Bei dickeren Schichten besteht außerdem die Möglichkeit, daß sich neben der kubischen Phase auch hexagonales CdS bildet [Zah94].

Stampfl et al. haben winkelaufgelöste Photoemissions-Messungen (ARPES) an kubischen CdS-Schichten auf InP(110) durchgeführt [Sta96]. Wir stellen in Tabelle 13.2 die gemessenen Daten unseren Quasiteilchen-Energien gegenüber. Bei den Cd $4d$ Zuständen beschränken wir uns auf die Schwerpunkte der fünf Bänder, da in den experimentellen Daten natürlich auch die Spin-Bahn Aufspaltung enthalten ist, in unseren Rechnungen hingegen nicht. Wir stellen in den oberen Valenzbändern eine gute Übereinstimmung unserer Ergebnisse mit den gemessenen Daten fest. Die bereits erwähnte Diskrepanz zwischen unseren berechneten Cd $4d$ Niveaus und den gemessenen Energien besteht auch mit dieser ARPES-Messung. Für das unterste S $3s$ artige Valenzband ergibt unsere Quasiteilchen-Rechnung eine Energie von –12.34

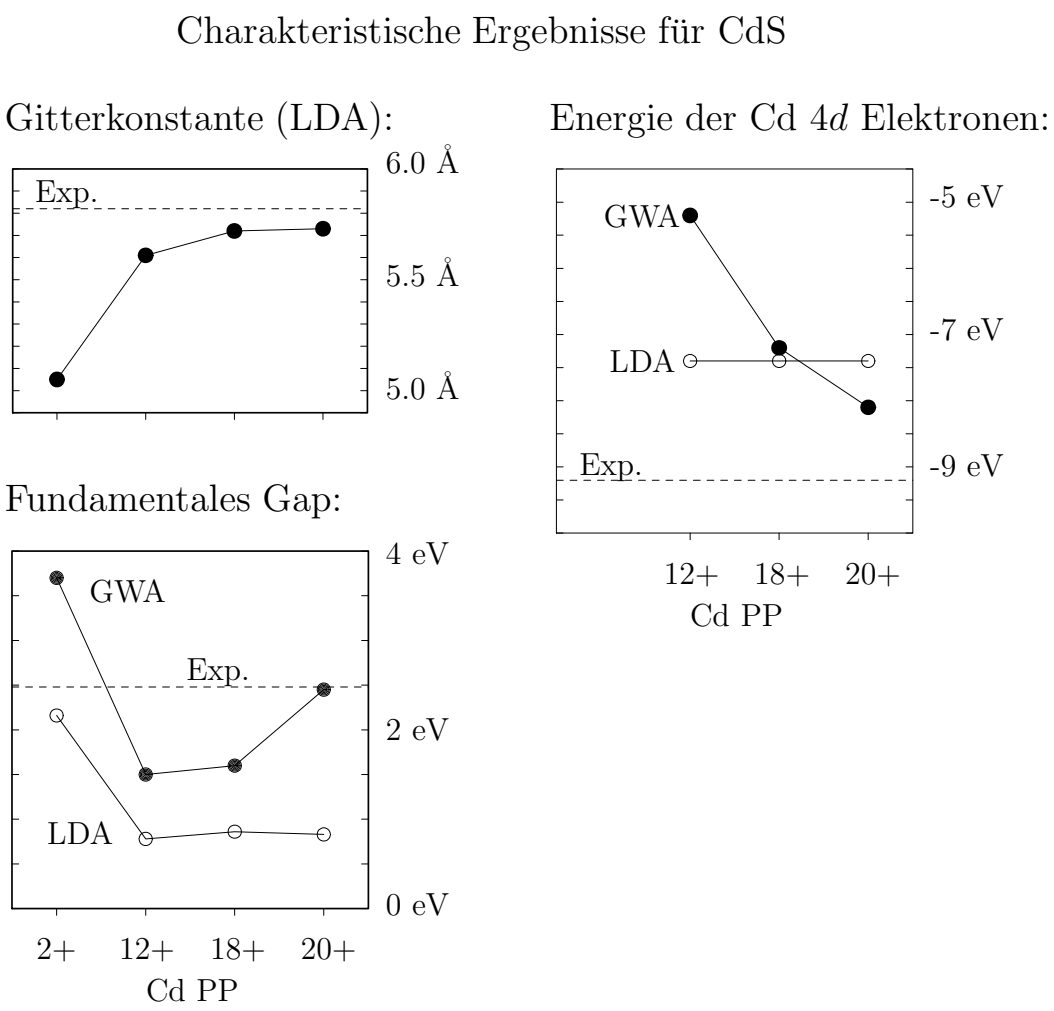


Abbildung 13.5: *Wesentliche Resultate für kubisches CdS, berechnet mit den vier verschiedenen Cd Pseudopotentialen.*

eV, während Stampfl et al. einen Wert von –13.85 eV messen. Diese Abweichung könnte auf allgemeine Schwierigkeiten bei der Ermittlung der Energie solch tiefer Valenzzustände sowohl im Experiment als auch in unserer Rechnung zurückzuführen sein. Wir erinnern daran, daß die verschiedenen Plasmon-Pol-Modelle bei den tieferen Valenzbändern zu leicht unterschiedlichen Quasiteilchen-Energien führen (vgl. Kap. 4).

Wir stellen in Abb. 13.5 nochmals die wesentlichen Veränderungen der berechneten Resultate zusammen, die sich ergeben, wenn die Cd Semicore Elektronen sukzessive in die Valenzelektronen überführt werden. Die theoretische LDA-Gitterkonstante wird monoten größer, wenn den Semicore-Elektronen die Möglichkeit

zur räumlichen Relaxation eingeräumt wird. Wenn ein Zustand im Kristall sich im Vergleich zum freien Atom räumlich nicht ändert, wie etwa der Cd 4*s* Zustand, so verändert sich auch die Gitterkonstante bei seiner Mitnahme als Valenzzustand praktisch nicht. Das endgültige Ergebnis für die Gitterkonstante stimmt gut mit dem experimentellen Wert überein. Die LDA-Bandstruktur ändert sich durch die Mitnahme der Cd 4*d* Elektronen als Valenzelektronen drastisch, während die Mitnahme der Cd 4*p* Elektronen nur noch geringe Auswirkungen hat. Durch die Mitnahme der Cd 4*s* Elektronen ändert sich die LDA-Bandstruktur überhaupt nicht. Für eine konvergente LDA-Behandlung ist es erforderlich und ausreichend, genau die Zustände als Valenzzustände mitzunehmen, die sich beim Übergang vom Atom zum Festkörper in ihrer räumlichen Struktur verändern. Im vorliegenden Falle genügt also für LDA-Rechnungen ein Cd^{18+} PP in jedem Fall, bei etwas geringerem Anspruch an die Rechengenauigkeit auch ein Cd^{12+} PP.

Bei den GW-Resultaten für die Quasiteilchen-Bandstruktur bietet sich ein völlig anderes Bild. Hier treten von Pseudopotential zu Pseudopotential deutliche Änderungen auf. Sie sind darauf zurückzuführen, daß von der LDA zur GWA der methodische Zugang wechselt, das Pseudopotential aber weiterhin auf der LDA beruht. Wir hatten zu Beginn des Kapitels das generelle Problem aufgezeigt, das sich daraus ergibt (siehe Gl. 13.4). Hier können wir nun auf der Basis unserer Ergebnisse die Größe dieser Fehler angeben. Sie betragen im vorliegenden Falle bis zu 4 eV. Konvergente Ergebnisse in guter Übereinstimmung mit experimentellen Daten lassen sich nur erzielen, indem die gesamte Semicore-Schale des Cd Atoms als Valenzschale berücksichtigt wird, also mit einem Cd^{20+} PP. Eine Unterteilung in Core- und Valenzelektronen führt zu unkontrollierbaren Unwägbarkeiten in den Ergebnissen, und die berechneten Bandstruktur-Energien haben dann keinerlei Aussagekraft.

Ein interessanter Aspekt betrifft die Quasiteilchen-*Korrekturen*, die mit den beiden Pseudopotentialen Cd^{2+} (gesamte Semicore-Schale im Core) bzw. Cd^{20+} (alle Semicore-Zustände als Valenzzustände) berechnet wurden. Bereits der Vergleich zwischen den zugehörigen Bandstrukturen (linke Seite in Abb. 13.1 und rechte Seite in Abb. 13.3) läßt erkennen, daß die Quasiteilchen-Korrekturen (= Differenz zwischen durchgezogenen Quasiteilchen-Bändern und gestrichelten LDA-Bändern) in beiden Fällen sehr ähnlich sind. Wir stellen in Abb. 13.6 diese Quasiteilchen-Korrekturen nochmals explizit dar, und zwar in Abhängigkeit von der LDA-Energie der jeweiligen Zustände. Volle Kreise markieren die Ergebnisse der Cd^{20+} PP Rechnung, offene Kreise die der Cd^{2+} PP Rechnung. Tatsächlich sind die beiden Datensätze

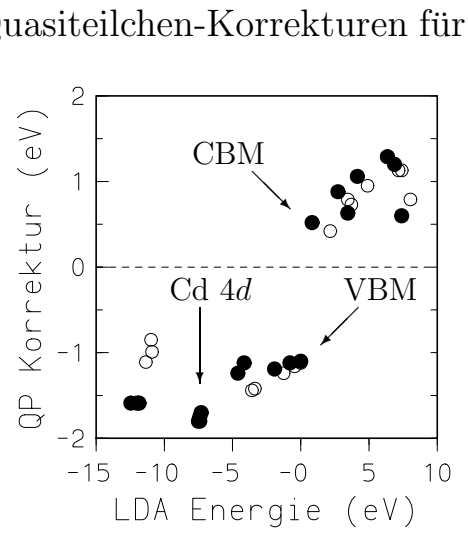


Abbildung 13.6: *Quasiteilchen-Korrekturen für CdS, berechnet mit dem Cd^{2+} PP (offene Kreise) sowie dem Cd^{20+} PP (volle Kreise).*

sehr ähnlich. Wesentliche Unterschiede ergeben sich nur beim untersten Valenzband (in Abb. 13.6 am linken Rand) und natürlich für die Cd 4*d* Elektronen, die in der Cd^{2+} PP Rechnung gar nicht behandelt werden. Offenbar lassen sich im Bereich des Energiegaps die endgültigen Quasiteilchen-Korrekturen der Cd^{20+} PP Rechnung bereits mit dem Cd^{2+} PP recht verläßlich bestimmen. Zur Berechnung von Quasiteilchen-Korrekturen ist es lediglich notwendig, daß alle Zustände einer Schale in gleicher Weise berücksichtigt werden – entweder alle als Core-Elektronen oder alle als Valenzelektronen. Die letztere Vorgehensweise liefert natürlich die zuverlässigeren Ergebnisse. Die starke Ähnlichkeit der beiden Datensätze ist auch der Grund, warum Zakharov et al. in ihrer Rechnung mit einem Cd^{2+} PP und nichtlinearen Core-Korrekturen Bandstruktur-Ergebnisse erzielen konnten, die mit unseren Resultaten recht gut übereinstimmen [Zak94].

13.2 Analyse der Selbstenergien in CdS

Im Folgenden soll genauer untersucht werden, warum das fundamentale Gap von kubischem CdS derartig stark auf die Mitnahme der Cd Semicore-Elektronen als Valenzelektronen reagiert. Wir stellen hierzu zunächst in Abb. 13.7 die Wellenfunktionen der VBM-Zustände und des CBM-Zustands dar, durch die das Gap definiert

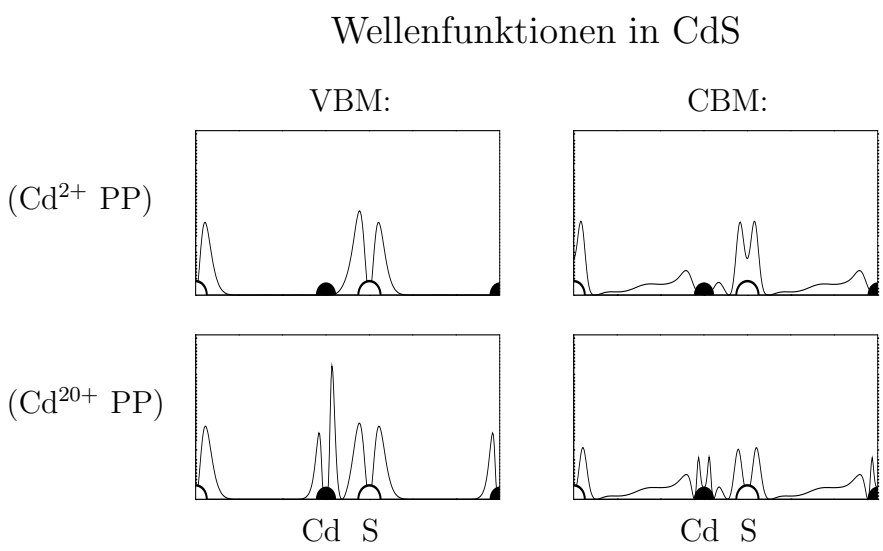


Abbildung 13.7: *Ladungsdichten der Wellenfunktionen des VBM und des CBM in kubischem CdS. Die Darstellung erfolgt entlang einer Cd-S Bindung ([111]-Richtung). Die fetten Punkte markieren Cd Atome, die offenen Kreise stellen S Atome dar. Die Wellenfunktionen in der oberen Zeile wurden mit dem Cd^{2+} PP berechnet, die in der unteren Zeile mit dem Cd^{20+} PP. Beim VBM wurde über die drei entarteten Zustände gemittelt.*

wird. In der Cd^{2+} PP Rechnung wird das VBM zum allergrößten Teil aus den S $3p$ Orbitalen gebildet. Die Berücksichtigung der Cd $4d$ Orbitale in den erweiterten Cd PP'en bewirkt eine Hybridisierung der VBM Zustände mit den $4d$ Orbitalen. Durch diese p-d Wechselwirkung erhält das VBM einen starken Anteil an p- und d-Orbitalen der Cd Atome. Der Zustand wird also durch die Erweiterung des Cd PP entscheidend verändert. Im Gegensatz dazu erfährt der CBM Zustand (rechte Seite in Abb. 13.7) keine starke Modifikation durch die Erweiterung des Pseudopotentials. Er besteht sowohl in der Cd^{2+} PP als auch in der Cd^{20+} PP Rechnung überwiegend aus den Cd $5s$ Orbitalen mit starken Beiträgen von s-Orbitalen der Schwefel-Atome. Eine wesentliche Änderung ergibt sich nur in unmittelbarer Nähe der Cd Kerne. Da in der Cd^{20+} PP Rechnung der Cd $4s$ Zustand explizit auftritt, muß der Cd $5s$ Zustand zu ihm orthogonal sein. Daher weist er in einem Abstand von 0.5 Å zum Cd Atomkern einen Knoten auf und erreicht bei noch kleineren Radien ein weiteres Extremum (in Abb. 13.7 (rechts unten) beiderseits des Cd Kerns zu erkennen).

Die beiden Zustände, die das fundamentale Gap definieren, sind also aus völlig unterschiedlichen Orbitalanteilen zusammengesetzt. Diese Zusammensetzung erfährt

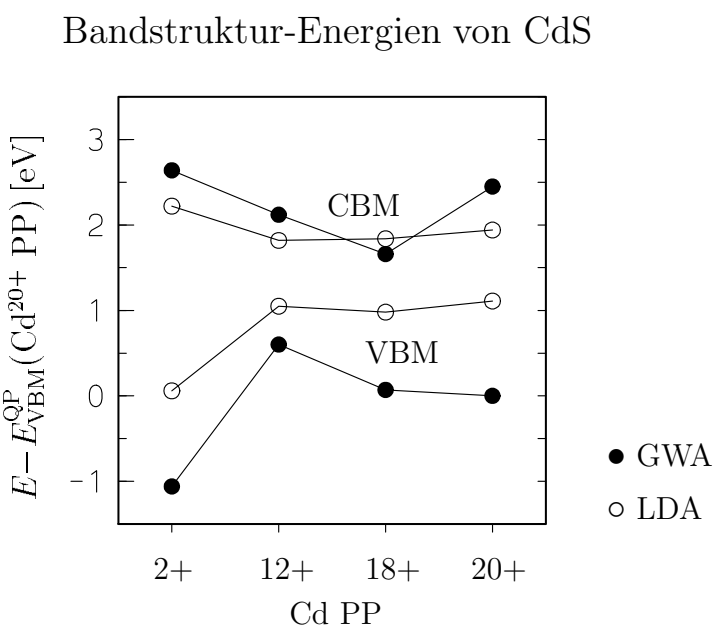


Abbildung 13.8: *Bandstruktur-Energien des VBM und des CBM von CdS auf einer absoluten Energieskala, die einen Vergleich der Rechnungen mit den vier verschiedenen Cd Pseudopotentialen erlaubt (siehe Text). Die Skala wurde so gewählt, daß die GW Quasiteilchen-Energie des VBM in der Cd^{20+} PP Rechnung Null beträgt. Fette Punkte (•) bedeuten GW Quasiteilchen-Energien, offene Kreise (∘) markieren LDA-Energien.*

bei dem einen Zustand (VBM) eine wesentliche Veränderung durch die Erweiterung des Cd Pseudopotentials, während der andere Zustand nur wenig verändert wird. Daher ist zu erwarten, daß die beiden Zustände in den verschiedenen Rechnungen auch sehr unterschiedliche Energien aufweisen werden. Das läßt sich deutlich erkennen, wenn alle Energiewerte für das VBM und das CBM in den LDA- und GWA-Rechnungen für die verschiedenen PP'e auf einer absoluten Energieskala miteinander verglichen werden. Dazu wird eine Referenzenergie benötigt, die nicht davon abhängt, welches Cd Pseudopotential in der jeweiligen Rechnung verwendet wird. Eine natürliche Vergleichsenergie wäre das Vakuumniveau über der Oberfläche des Festkörpers. Die Durchführung der dazu notwendigen Slabrechnungen erweist sich wegen der Semicore-Elektronen als sehr aufwendig und wird von uns nicht verfolgt. Statt dessen wählen wir als gemeinsame Referenzenergie den LDA-Potentialwert im interstitiellen Bereich. Der maximal erreichbare Abstand eines Punktes im Kristall zu den nächsten Cd Atomen beträgt $a_0/2 = 2.91$ Å $= 5.50$ a.u. Bei diesem Abstand wird der Verlauf der Potentiale nur noch unwesentlich davon beeinflußt, welches

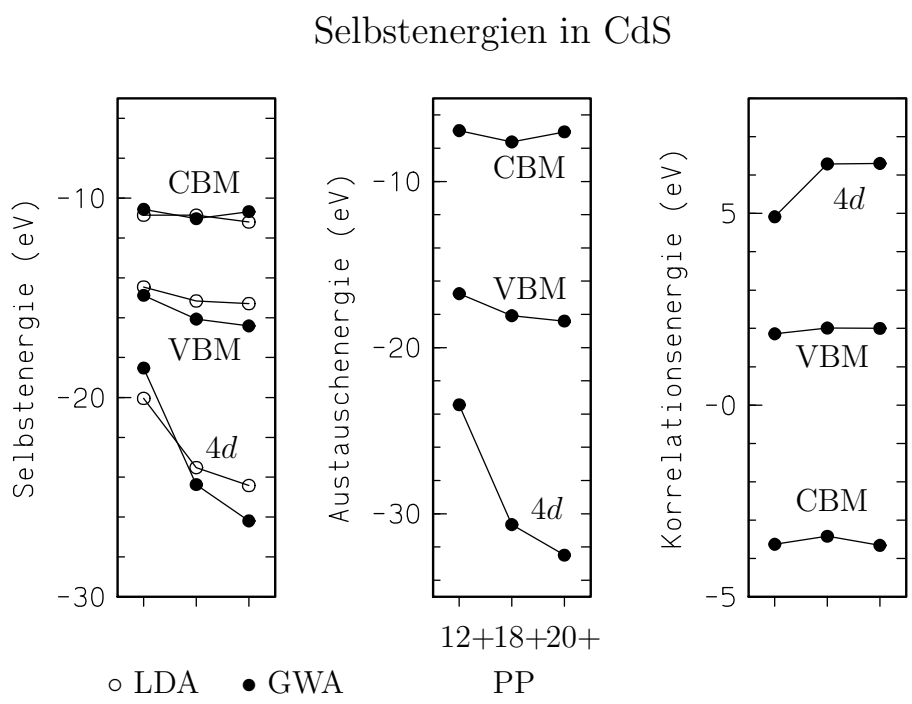


Abbildung 13.9: *Berechnete Selbstenergiewerte für elektronische Zustände in CdS. Linke Figur: gesamte GW-Selbstenergie (•) und LDA Austausch-Korrelations-Potential (○). Mittlere/rechte Figur: Austausch- und Korrelations-Beiträge zur GW-Selbstenergie. Es werden Rechnungen mit drei verschiedenen Cd Pseudopotentialen miteinander verglichen.*

der vier Cd Pseudopotentiale verwendet wird. Wir beziehen alle GWA und LDA Bandstruktur-Ergebnisse auf den lokalen LDA Potentialwert an diesem Punkt. Das ermöglicht einen direkten Vergleich zwischen den Rechnungen mit den verschiedenen Pseudopotentialen. Die resultieren Energiewerte stellen wir in Abb. 13.8 dar. Da der Potentialwert an dem interstitiellen Punkt keine eigene Bedeutung besitzt, haben wir die gesamte Energieskala so verschoben, daß das GW Quasiteilchen-VBM der Cd^{20+} Rechnung den Wert Null erhält.

Bei der Hinzunahme der Cd $4d$ Elektronen in das Cd^{12+} PP erfährt das p-artige VBM eine starke Anhebung (1 eV in LDA, 1.5 eV in GWA) durch die nun mögliche p-d Wechselwirkung mit den Cd $4d$ Zuständen. Die gleichzeitige Absenkung des CBM fällt wesentlich kleiner aus. Beide Effekte zusammen reduzieren das LDA-Gap um 1.33 eV und das GWA-Gap um 2.20 eV. Die Berücksichtigung der weiteren Cd Semicore-Elektronen im Cd^{18+} und Cd^{20+} PP beeinflußt die LDA-Energien erwartungsgemäß kaum. Auf die GW Quasiteilchen-Energien haben die $4s$ und $4p$ Elektronen hingegen einen erheblichen Einfluß. Durch die $4p$ Elektronen im PP wird die

p-d Wechselwirkung zwischen den Cd 4*d* Zuständen und dem VBM im Rahmen der GWA wesentlich besser beschrieben, und die GW-Energie des VBM wird um 0.53 eV abgesenkt. Gleichzeitig wird aber auch das CBM abgesenkt, so daß das GW-Energiegap sich von der Cd^{12+} PP zur Cd^{18+} PP Rechnung praktisch nicht ändert. Erst die Mitnahme der Cd 4*s* Elektronen im Cd^{20+} PP führt zu einer Anhebung des CBM, das zum großen Teil aus dem Cd 5*s* Zustand besteht. Dadurch vergrößert sich auch das fundamentale Gap auf 2.45 eV.

Die systematische Differenz $\Delta\Sigma_{\rm xc}$ (13.4) zwischen LDA und GWA verdient ebenfalls noch eine genauere Untersuchung. Wir stellen in Abb. 13.9 dar, wie sich die Selbstenergien der einzelnen Zustände und ihre Beiträge ändern, wenn das Pseudopotential systematisch erweitert wird. Dabei konzentrieren wir uns auf die Cd 4*d* Elektronen, auf das VBM und das CBM. Bei allen Zuständen treten die stärksten Veränderungen beim Austauschanteil auf (mittlere Figur), während der Korrelations-Anteil (rechte Figur) nur schwach variiert. Die gesamte Selbstenergie (linke Figur) der Cd 4*d* Elektronen ändert sich durch die Mitnahme der Cd 4*s* und 4*p* Elektronen sehr stark (um 8 eV). Bei den anderen Zuständen (VBM, CBM) fällt die Änderung wesentlich geringer aus (ca. 1 eV). Die LDA liefert geringere Veränderungen der Selbstenergie (man beachte die unterschiedlichen Steigungen der LDA- und der GWA-Kurven). Daraus resultieren die unterschiedlichen Quasiteilchen-Korrekturen für die verschiedenen Cd Pseudopotentiale Cd^{12+}, Cd^{18+} und Cd^{12+}. [Wir erinnern daran, daß sich durch die Berücksichtigung der Cd 4*s* und 4*p* Zustände die LDA Bandstruktur *nicht* ändert – wohl aber die Unterteilung der XC-Energie in den Pseudopotentialanteil und den in Abb. 13.9 dargestellten expliziten Valenzanteil. Die Summe beider Terme bleibt nahezu gleich.]

Wir wollen noch untersuchen, an welcher Stelle des GW-Formalismus die Semicore-Elektronen die GW Selbstenergie wesentlich beeinflussen. Für den Einfluß kommen hauptsächlich die folgenden beiden Mechanismen in Betracht.

(i) Die Semicore-Elektronen besitzen eine (wenn auch geringe) Polarisierbarkeit [Shi92, Shi93b]. Bei der Mitnahme der Elektronen als Valenzelektronen wird sie automatisch berücksichtigt und beeinflußt die abgeschirmte Wechselwirkung und schließlich den Korrelationsbeitrag zur Selbstenergie. Eine systematische Analyse dieses Einflusses läßt sich durchführen, indem die Semicore-Elektronen bei der Berechnung der Polarisierbarkeit gemäß Gl. (3.12) gezielt ignoriert werden. Wir stellen in Tabelle 13.3 die Ergebnisse einer solchen Untersuchung vor. Alle Rechnungen wurden im Rahmen der Cd^{20+} PP Rechnung durchgeführt. In der ersten Spalte wurden

Einfluß der Semicore-Zustände auf die Polarisierbarkeit

	P ohne $4s$, p, d	P ohne $4s$, p	volles P
berücksichtigt:			
Cd $4s$, $4p$			×
Cd $4d$		×	×
Rest	×	×	×
ϵ_∞	6.22	6.43	6.44
	Korrelations-Energie [eV]		
Cd $4d$	5.17	6.14	6.30
VBM	1.97	2.02	2.00
CBM	–3.27	–3.58	–3.66

Tabelle 13.3: *Berechnete makroskopische dielektrische Funktion und berechnete Korrelations-Beiträge zur Selbstenergie verschiedener Zustände am Γ-Punkt für CdS. Bei der Bestimmung der Polarisierbarkeit P im Rahmen der RPA wurden die Cd Semicore-Elektronen teilweise unterdrückt. Alle Berechnungen erfolgen im Rahmen der Cd^{20+} PP Rechnung.*

in der Polarisierbarkeit alle Cd Semicore-Elektronen ($4s$, $4p$, $4d$) unterdrückt. Für die zweite Spalte wurden nur die tiefliegenden Semicore-Elektronen ($4s$, $4p$) unterdrückt, die Cd $4d$ Elektronen hingegen berücksichtigt. Die letzte Spalte enthält die Ergebnisse der gewöhnlichen Cd^{20+} Rechnung, in der alle Beiträge zur Polarisierbarkeit berücksichtigt wurden.

Für die makroskopischen Abschirmeffekte sind die Semicore-Elektronen nur von geringer Bedeutung. Die dielektrische Konstante ändert sich bei der Berücksichtigung der Polarisierbarkeit der Cd $4d$ Elektronen nur um 3 % (von 6.22 auf 6.43). Die Hinzunahme der Cd $4s$ und $4p$ Elektronen führt praktisch zu keiner weiteren Veränderung. Unser Ergebnis von 6.44 für die dielektrische Konstante ist etwas größer als der experimentelle Wert von 5.3 [Lan82]. Der Korrelationsbeitrag zur Selbstenergie der einzelnen Zustände hängt nur sehr wenig davon ab, ob die Semicore-Elektronen in der Polarisierbarkeit berücksichtigt werden oder nicht. Lediglich die Korrelationsenergie der Cd $4d$ Elektronen erfährt eine Anhebung um 1eV durch Einbeziehung ihrer eigenen Polarisierbarkeit in die Abschirmung. Die geringe Bedeutung der Semicore-Elektronen für die dielektrische Funktion erklärt sich aus ihrer starken Lokalisierung und bei den Cd $4s$ und $4p$ Elektronen zusätzlich aus ihrer

Zusammensetzung der Selbstenergien

Selbstenergie des Zustands		Anteile [eV] von Cd 4s, p	Cd 4d	Rest
Cd 4d	Σ_x	–9.39	–19.70	–1.91
	Σ_c	0.49	5.31	–0.63
VBM	Σ_x	–1.56	–2.67	–12.08
	Σ_c	0.08	0.57	0.28
CBM	Σ_x	–0.70	–1.46	–4.86
	Σ_c	–0.07	–0.21	–2.99

Tabelle 13.4: *Additive Beiträge verschiedener Gruppen von Zuständen zur Selbstenergie ausgewählter Zustände in CdS. Alle Werte beziehen sich auf die Cd^{20+} PP Rechnung. Mit „Rest“ fassen wir alle anderen Zustände außer den Cd Semicore-Zuständen zusammen. Die Berechnung der Korrelationsenergien der einzelnen Zustände erfolgte an der jeweiligen LDA-Energie.*

großen Bindungsenergie, die in die Polarisierbarkeit als Nenner eingeht.

(ii) Über die Green-Funktion G beeinflussen die Semicore-Elektronen direkt den Selbstenergie-Operator $\Sigma = GW$. Die Selbstenergie eines jeden Zustands setzt sich additiv aus einzelnen Beiträgen von allen anderen Zuständen zusammen (vgl. Gl.(4.12)). Wir haben in unserer Cd^{20+} PP Rechnung die Selbstenergien nach diesen Beiträgen aufgeschlüsselt. Die Ergebnisse sind in Tabelle 13.4 angegeben. Die Selbstenergie der Cd 4d Zustände beinhaltet einen äußerst starken Austausch-Beitrag von den Cd 4s und Cd 4p Zuständen, der etwa ein Drittel der gesamten Austausch-Energie ausmacht. Darin zeigt sich nochmals die essentielle Bedeutung dieser Zustände für eine quantitativ richtige Beschreibung der Cd 4d Energie. Der Anteil dieser Zustände zur Korrelationsenergie von Cd 4d ist weit weniger bedeutsam; er macht nur etwa 10 % aus. Das liegt an dem großen Energieunterschied zwischen den zugehörigen Bändern, der als Nenner in die Korrelationsenergie eingeht. Im Gegensatz zu Cd 4d beinhaltet die Selbstenergie der VBM- und CBM-Zustände nur geringe Beiträge der Cd 4s und 4p Zustände. Ihre Beiträge zur Korrelationsenergie liegen hier unter 0.1 eV.

Erwartungsgemäß stammen die größten Beiträge zur Selbstenergie der Cd 4d Zustände von den Zuständen selbst. Die restlichen Valenzband- und Leitungsband-Zustände tragen hierzu relativ wenig bei. Umgekehrt tragen die Cd 4d Zustände

Pseudopotentiale für Cd, Zn und Ga

	ohne Semicore		mit Semicore	
	PP	Zustände	PP	Zustände
Cd	2+	$5s$, $5p$	20+	$4s$, $4p$, $4d$; $5s$
Zn	2+	$4s$, $4p$	20+	$3s$, $3p$, $3d$; $4s$
Ga	3+	$4s$, $4p$	21+	$3s$, $3p$, $3d$; $4s$, $4p$

Tabelle 13.5: *Struktur der Pseudopotentiale für Zn, Cd und Ga. Wir geben an, welche Elektronen jeweils als Valenzelektronen behandelt werden.*

ihrerseits nur wenig zu den Selbstenergien der VBM- und CBM-Zustände bei. Allerdings sind die Beiträge von Cd $4s$, $4p$ und $4d$ zur Selbstenergie des VBM größer als zur Selbstenergie des CBM. Daraus resultiert die schon mehrfach erwähnte Abhängigkeit des fundamentalen Gaps von dem Cd Pseudopotential der jeweiligen Rechnung.

13.3 Weitere Verbindungshalbleiter

In diesem Abschnitt diskutieren wir die Bandstrukturen von weiteren Verbindungshalbleitern, in denen kationische Semicore-Zustände eine Rolle spielen. Wir behandeln exemplarisch die kubischen Modifikationen von ZnS und GaN.

Als zentrales Ergebnis des vorletzten Abschnitts hatten wir festgestellt, daß zur zuverlässigen Berechnung von Quasiteilchen-Korrekturen alle Zustände einer jeden Schale bei der Pseudopotential-Konstruktion gleich behandelt werden müssen. Die Zerlegung einer Schale in Valenz- und Core-Elektronen führt zu unkontrollierbaren Fehlern in den Quasiteilchen-Korrekturen und in der Quasiteilchen-Bandstruktur. Daher beschränken wir uns bei allen weiteren Untersuchungen auf zwei Typen von Rechnungen. Wir führen einerseits Rechnungen mit solchen Pseudopotentialen durch, die nur die konventionellen Valenzelektronen der unvollständig besetzten Valenzschale beschreiben [Bac82]. Dieses sind die Pseudopotentiale Zn^{2+} bzw. Ga^{3+}. Weiterhin behandeln wir die Materialien auch mit Pseudopotentialen, für die wir die gesamte Semicore-Schale als Valenzschale mit berücksichtigen. Das führt zu den Pseudopotentialen Zn^{20+} und Ga^{21+}. Wir stellen die einzelnen Pseudopotentiale in Tabelle 13.5 zusammen.

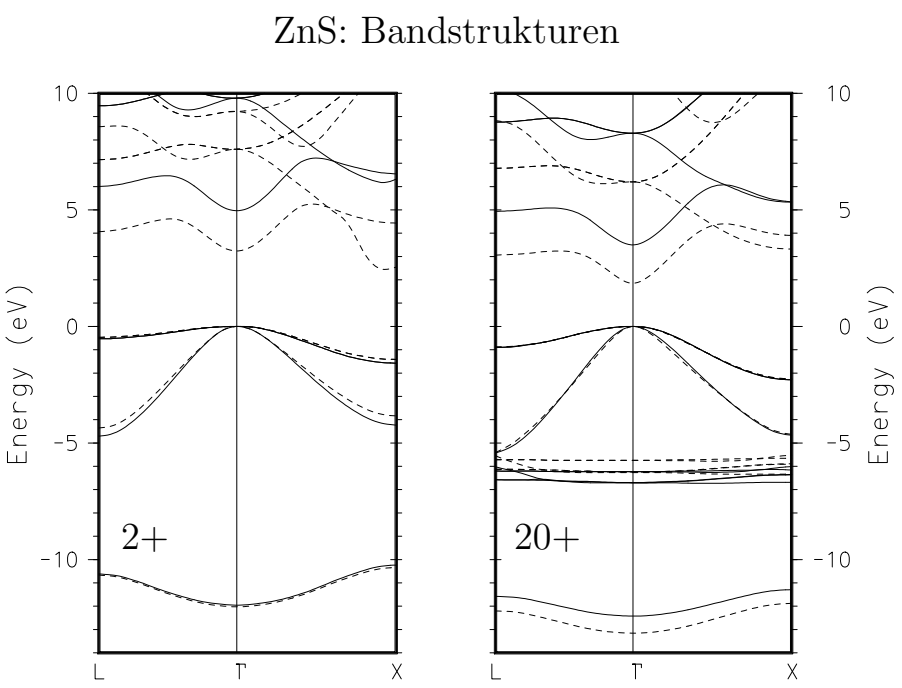


Abbildung 13.10: *Berechnete Bandstrukturen für kubisches ZnS. Durchgezogene Kurven: GWA Quasiteilchen-Bandstruktur. Gestrichelt: LDA-Energien. Es werden die Ergebnisse dargestellt, die mit dem Zn^{2+} PP (links) sowie mit dem Zn^{20+} PP (rechts) berechnet wurden.*

Die kubische Modifikation von ZnS (Zinkblende) wurde theoretisch bereits intensiv untersucht [Eng90, Sch93a, Zak94]. In der Studie von Schröer et al. [Sch93a] wurde nachgewiesen, daß die Berücksichtigung der Zn 3*d* Zustände als Valenzzustände notwendig ist, um zu einer konsistenten Beschreibung der strukturellen und elektronischen Eigenschaften zu gelangen.

Zn weist die gleiche Elektronenkonfiguration wie Cd auf. Es besitzt lediglich eine Elektronenschale weniger. Bei kubischem ZnS liegen daher die Zn 3*d* Zustände ähnlich wie bei CdS innerhalb der Valenzbänder. Ihre LDA-Energie beträgt im Mittel –5.9 eV unterhalb des VBM. Der Kopplungsmechanismus dieser Zustände mit den S 3*p*-artigen oberen Valenzbändern verläuft völlig analog zu demjenigen, den wir im letzten Abschnitt bei CdS diskutiert hatten. Wir untersuchen hier lediglich einige Aspekte von ZnS, um quantitative Vergleiche zu CdS vornehmen zu können. In Abb. 13.10 stellen wir die GW Quasiteilchen-Bandstrukturen und die LDA-Bandstrukturen (gestrichelt) vor, die wir mit dem Zn^{2+} PP (links) sowie dem Zn^{20+} PP (rechts) berechnet haben.

Die Bandstrukturen ähneln sehr stark den entsprechenden Bandstrukturen für CdS. Die Behandlung des Materials mit dem Zn^{2+} PP führt auf eine Quasiteilchen-Bandstruktur, die deutliche Abweichungen von den experimentellen Daten aufweist. Das Quasiteilchen-Gap ist mit 4.97 eV weit größer als das gemessene Gap von 3.8 eV [Lan82]. Die Bandbreite der oberen Valenzbänder beträgt 4.70 eV in unserer Quasiteilchen-Bandstruktur; sie ist im Vergleich zum experimentellen Wert (5.5 eV [Lan82]) zu gering. In der LDA-Bandstruktur nimmt das unterste Leitungsband sein Minimum nicht am Γ-Punkt, sondern am X-Punkt an. Das widerspricht der experimentellen Beobachtung, daß ZnS ein direkter Halbleiter ist. Immerhin wird diese falsche Dispersion durch die Quasiteilchen-Korrekturen richtiggestellt, denn das fragliche Band erfährt am X-Punkt eine besonders große Energieanhebung von +3.78 eV.

Die Rechnung mit dem Zn^{20+} PP liefern eine Quasiteilchen-Bandstruktur, die mit den experimentellen Daten weit besser übereinstimmt. In dieser Rechnung treten die Zn $3d$ Elektronen explizit auf. Ihre LDA-Energie von –5.9 eV ist im Vergleich zum experimentellen Wert von –9.0 eV [Lan82] zu hoch. Durch die Quasiteilchen-Korrekturen werden ihre Energien um –0.5 eV abgesenkt; ihre Quasiteilchen-Energien betragen –6.4 eV. Auch dieser Wert liegt höher als das experimentelle Resultat. Dieses Ergebnis ist völlig analog zu der Unterschätzung der Cd $4d$ Quasiteilchen-Bindungsenergie in CdS zu verstehen. Wir werden in Kap. 14 einen erweiterten Zugang zu den Bindungsenergien stark lokalisierten Zustände vorstellen. Durch die Berücksichtigung der Zn Semicore Zustände im Zn^{20+} PP erfährt das LDA-Gap von ZnS eine Absenkung von 1.38 eV gegenüber der Zn^{2+} PP Rechnung. Die Ursache liegt wie bei CdS im Auftreten der p-d Wechselwirkung zwischen den Zn $3d$ Zuständen und dem VBM. Das Quasiteilchengap wird in gleicher Weise um 1.47 eV gegenüber der Zn^{2+} PP Rechnung reduziert. Es beträgt nun 3.50 eV und stimmt mit dem experimentellen Wert von 3.8 eV recht gut überein. Die Bandbreite der oberen Valenzbänder war in der Zn^{2+} PP Rechnung unterschätzt worden. In der Zn^{20+} PP Rechnung erhalten wir nun eine Quasiteilchen-Bandbreite von 5.42 eV (exp.: 5.5 eV).

In Abb. 13.11 stellen wir für CdS, ZnS und GaN die mit den verschiedenen PP'en berechnetenen Quasiteilchen-Bandstrukturen in der Nähe des fundamentalen Gaps im direkten Vergleich dar. Die Veränderungen in den Quasiteilchen-Bandstrukturen durch die Hinzunahme der Semicore-Zustände fallen in CdS und in ZnS völlig gleichartig aus. Allerdings wird in ZnS das Gap etwas stärker reduziert (–1.47 eV) als

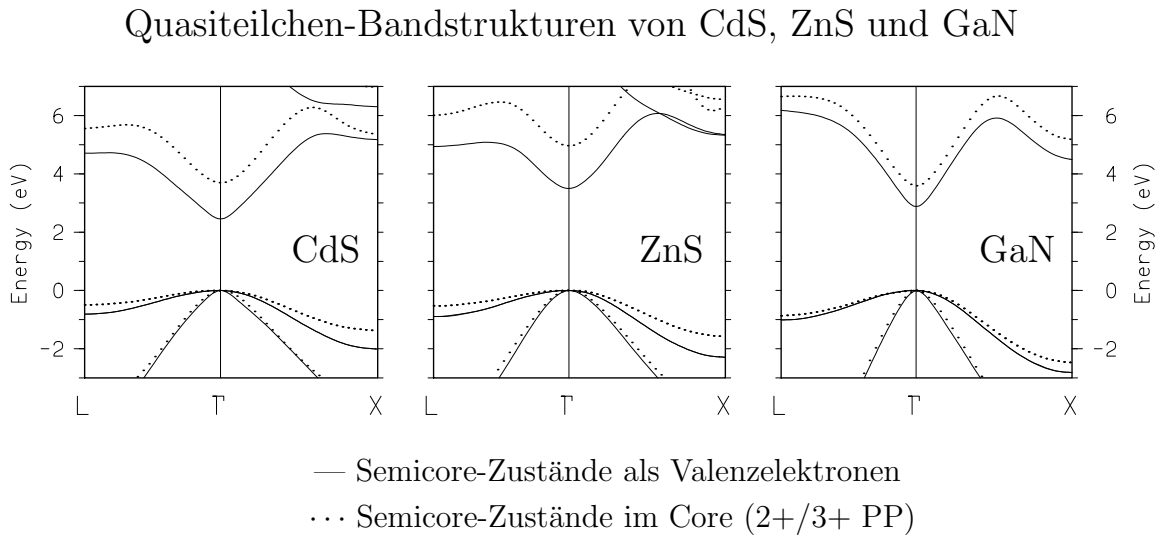


Abbildung 13.11: *GW Quasiteilchen-Bandstrukturen für die kubischen Modifikationen von CdS (links), ZnS (Mitte) und GaN (rechts). Wir vergleichen Bandstrukturen, zu deren Berechnung die Semicore-Zustände explizit berücksichtigt wurden (durchgezogene Kurven), mit Rechnungen, die mit 2+ bzw. 3+ Pseudopotentialen durchgeführt wurden (gepunktet; Semicore-Elektronen im Core).*

in CdS (–1.25 eV). Das war bereits bei den LDA-Bandstrukturen der Fall (Gap-Absenkung: –1.38 eV für ZnS, –1.33 eV für CdS). Die Analogie der Veränderungen der Bandstruktur der beiden Materialien ist auf die gleichartige orbitale Zusammensetzung der Wellenfunktionen und auf die gleichartige *p*-*d* Wechselwirkung zurückzuführen.

Wir wenden uns nun dem III-V Halbleiter GaN zu, einem Material, das ein enormes Anwendungspotential bei der Herstellung optoelektronischer Bauelemente im Bereich blauen Lichtes besitzt. Ebenso wie die II-VI Verbindungen ist GaN ein direkter Halbleiter mit einer Gapenergie von ca. 3.2 eV [Lei92] in der kubischen und 3.5 eV [Lan82] in der hexagonalen Modifikation. Die kubische Modifikation kann durch Aufbringen auf ein kubisches Substrat, etwa GaAs(100) oder GaAs(110), hergestellt werden [Din96].

Kubisches GaN wurde unter anderem von Fiorentini et al. [Fio93] sowie Wright und Nelson [Wri94] theoretisch untersucht. In beiden Arbeiten wird die Bedeutung der Ga 3*d*-Elektronen für eine konsistente Beschreibung des Materials betont. Rubio et al. [Rub93] sowie Palumno et al. [Pal94] haben die Quasiteilchen-Bandstruktur von GaN berechnet. In beiden Arbeiten werden allerdings die *d* Elektronen nicht

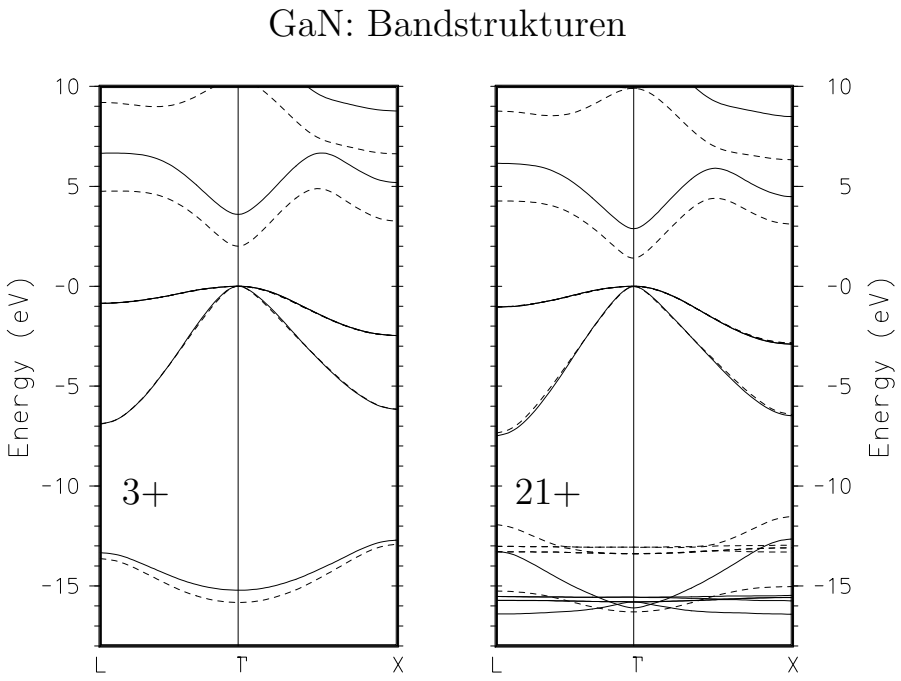


Abbildung 13.12: *Berechnete Bandstrukturen für kubisches GaN. Durchgezogene Kurven: GWA Quasiteilchen-Bandstruktur. Gestrichelt: LDA-Energien. Es werden die Ergebnisse dargestellt, die mit dem Ga^{3+} PP (links) sowie mit dem Ga^{21+} PP (rechts) berechnet wurden.*

explizit behandelt, sondern es wird ein Ga^{3+} PP unter Berücksichtigung von Core-Korrekturen verwendet.

Ga besitzt eine um ein Proton größere Kernladung und ein Valenzelektron mehr als Zn. Daher ist die atomare Wellenfunktion des 3*d* Zustandes in Ga etwas stärker lokalisiert als in Zn. In Zn beträgt der mittlere Radius des 3*d* Zustands 0.48 Å, in Ga hingegen nur 0.43 Å. Entsprechend ist die atomare LDA Energie des Zustands in Ga mit –19.40 eV niedriger als in Zn (–10.42 eV). Wir erwarten daher, daß die Einflüsse dieses Zustandes und der anderen Zustände der dritten Schale auf die Bandstruktur von GaN geringer ausfallen als bei ZnS. Um das zu prüfen, haben wir wiederum Rechnungen mit zwei Ga Pseudopotentialen (Ga^{3+} PP und Ga^{21+} PP) durchgeführt. Wir stellen in Abb. 13.12 die resultierenden Bandstrukturen vor. Durchgezogene Linien markieren GW-Quasiteilchen-Bänder, gestrichelte Linien bedeuten LDA-Bänder.

Wir beschränken uns in der Diskussion auf zwei Aspekte der Bandstrukturen.
(i) Das fundamentale Gap ergibt sich in der Ga^{3+} PP Rechnung zu 2.00 eV in LDA

und zu 3.59 eV in der GW Quasiteilchen-Bandstruktur (linke Figur in Abb. 13.12). Das Quasiteilchen-Gap ist im Vergleich zum Experiment (3.2 eV) zu groß. Nach den Ergebnissen für CdS und ZnS ist das kein überraschendes Resultat. Allerdings fällt die Überschätzung des Gaps bei GaN wesentlich geringer aus. Darin zeigt sich, daß der Einfluß der kationischen Semicore-Zustände bei Elementen aus höheren Gruppen schwächer ist als bei Elementen der Gruppe II. Dementsprechend wird durch die Hinzunahme der Ga Semicore-Zustände zu den Valenzzuständen das Quasiteilchengap nur um 0.71 eV reduziert (vgl. auch Abb. 13.11). Unser resultierendes Quasiteilchen-Gap von 2.88 eV liegt ein wenig unter dem experimentellen Wert von 3.2 eV.
(ii) Die unteren Valenzbänder verdienen besondere Beachtung. In der linken Figur in Abb. 13.12 findet sich ein einzelnes Band bei Energien zwischen –16 eV und –13 eV. Es wird von den N 2s Orbitalen gebildet. Seine Bandbreite beträgt in der vorliegenden Ga^{3+} PP Rechnung sowohl in LDA als auch in GWA etwa 3 eV (siehe auch [Rub93, Wri94]). In der Ga^{21+} PP Rechnung (rechte Figur in Abb. 13.12) werden die d Zustände explizit als Valenzelektronen berücksichtigt. Sie haben in GaN tiefere Bandstrukturenergien (–13 eV in LDA; gestrichelte Bänder) als in den II-VI Halbleitern. Dementsprechend koppeln sie nicht an die oberen Valenzbänder, sondern an das unterste, N 2s-artige Band. Am Rand der Brillouinzone (an L und an X) besitzt das oberste LDA-Band der ansonsten fast entarteten Ga 3d Zustände einen starken Beitrag von den N 2s Bändern und liegt um 1.5 eV über den anderen d Bändern. Entsprechend beinhaltet das N 2s Band erhebliche Beiträge von den Ga 3d Zuständen. Dadurch wird seine LDA-Bandbreite stark reduziert. Sie beträgt nunmehr 1.3 eV (0.8 eV in [Fio93], 1.3 eV in [Wri94]).

Wir haben zunächst auf der Basis dieser LDA-Rechnung den Selbstenergie-Operator im Rahmen der GWA konstruiert und die Störmatrix (2.13) untersucht. Der hauptsächliche Effekt dieser Störung besteht in der energetischen Absenkung der Ga 3d Orbitale um etwa 3 eV. Dadurch ändert sich die s-d-Kopplung mit den N 2s Orbitalen erheblich, und die Struktur, Energien und orbitale Zusammensetzung der Zustände erfahren starke Modifikationen. Das führt dazu, daß der Selbstenergie-Operator auf der Grundlage der LDA-Größen nicht selbstkonsistent berechnet werden kann. Die Genauigkeit des vorliegenden Selbstenergie-Operators reicht daher für eine quantitative Bestimmung der Quasiteilchen-Bandstruktur der Ga 3d und N 2s Zustände nicht aus (auch nicht im Rahmen einer expliziten Neu-Diagonalisierung, wie sie im nächsten Abschnitt im Falle der Bandstruktur von Ge diskutiert wird). Vielmehr erweist es sich als notwendig, schon bei der Konstruktion des Selbstenergie-

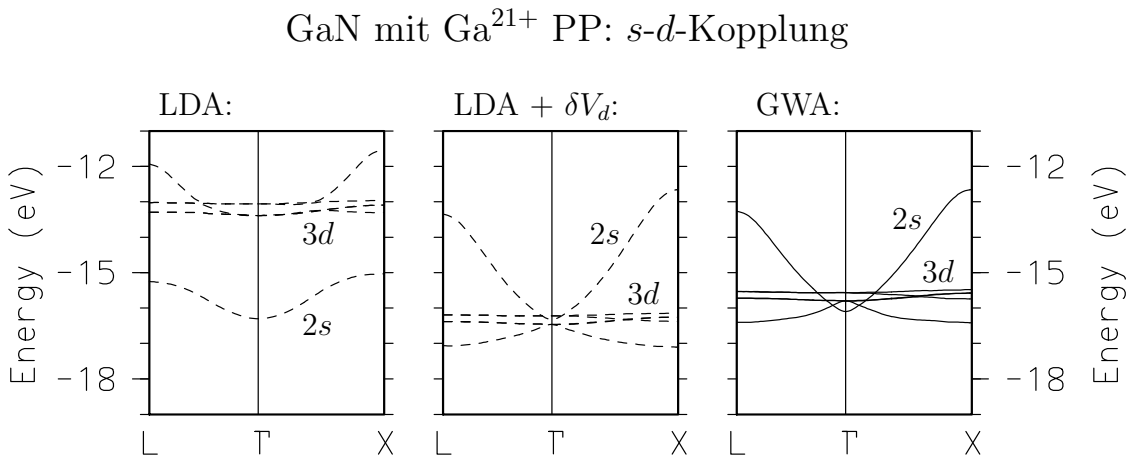


Abbildung 13.13: *Struktur der unteren Valenzbänder Ga 3d und N 2s in GaN im Rahmen der Ga^{21+} PP Rechnungen (siehe Text).*

Operators die tiefere energetische Lage der Ga $3d$ Zustände sowie die gegenüber der LDA geänderte Kopplung mit den N $2s$ Zuständen zu berücksichtigen. Wir erzielen das, indem wir im Rahmen der LDA ein zusätzliches, an den Ga Atomen lokalisiertes Hilfspotential δV_d berücksichtigen, das im wesentlichen nur auf die Ga $3d$ Orbitale wirkt und ihre Energie absenkt. Die s-d Kopplung ändert sich durch die niedrigere Energie der Ga $3d$ Orbitale automatisch. Wir stellen in Abb. 13.13 die resultierende Dispersion der Bänder dar (mittlere Figur). Es stellt sich heraus, daß im Rahmen dieser Rechnung die Dispersion des N $2s$ Bandes erheblich größer ist als in der originalen LDA-Rechnung (linke Figur). Das Band weist nun fast wieder den Verlauf auf, der sich im Rahmen der Ga^{3+} PP LDA-Rechnung ergeben hatte (vgl. Abb. 13.12, links). Die Ga $3d$ Bänder liegen in der vorliegenden Bandstruktur im Bereich des Minimums des N $2s$ Bandes (wir haben in Hinblick auf die noch zu diskutierende Quasiteilchen-Bandstruktur das zusätzliche Potential δV_d dementsprechend gewählt). Das unterste der fünf Ga $3d$ Bänder weist eine stärkere Dispersion als die restlichen vier Bänder auf. Der zugehörige Zustand enthält Anteile der N $2s$ Orbitale.

Auf der Grundlage dieser Bandstruktur konstruieren wir den Selbstenergie-Operator im Rahmen der GWA und bestimmen die Quasiteilchen-Korrekturen und die Quasiteilchen-Bandstruktur. Dabei ist Folgendes zu beachten. Im Rahmen der LDA stellt sich der Hamilton-Operator als Summe zweier Terme dar: $H^{\mathrm{LDA}} = H^0 + V_{\mathrm{xc}}$, wobei H^0 den selbstenergiefreien Anteil und V_{xc} die Selbstenergie darstellt. Im

Rahmen der GWA wird der zweite Term durch den GW Selbstenergie-Operator ersetzt:

$$\begin{aligned} H^{\mathrm{GWA}} &= H^0 + \Sigma \\ &= H^{\mathrm{LDA}} + (\Sigma - V_{\mathrm{xc}}) \quad . \end{aligned} \tag{13.6}$$

Wenn im Rahmen der LDA-Rechnung das zusätzliche Hilfspotential δV_d addiert wird, so berechnet sich der GW Hamilton-Operator wie folgt:

$$H^{\mathrm{GWA}} = (H^{\mathrm{LDA}} + \delta V_d) + (\Sigma - V_{\mathrm{xc}} - \delta V_d) \quad . \tag{13.7}$$

Das zusätzliche Potential muß also im Rahmen des Störoperators wieder subtrahiert werden. Wir messen dem Potential an dieser Stelle keine eigene physikalische Bedeutung zu, sondern verstehen es als künstliche Hilfsgröße zur möglichst einfachen Bestimmung eines selbstkonsistenten Selbstenergie-Operators. Die resultierende Quasiteilchen-Bandstruktur ist in Abb. 13.13 (rechts) dargestellt. Es stellt sich heraus, daß die Struktur der Bänder sich gegenüber der zugrunde liegenden Bandstruktur (mittlere Figur) praktisch nicht ändert. Dadurch ist die Berechnung des Selbstenergie-Operators im Rahmen der um δV_d modifizierten LDA gerechfertigt.

Wir vergleichen nun die resultierende GW Quasiteilchen-Bandstruktur mit den ursprünglichen LDA-Resultaten (siehe Abb. 13.12). In der LDA-Rechnung lagen die Ga 3*d* Bänder über dem N 2*s* Band. Durch die Quasiteilchen-Korrekturen wird diese Reihenfolge der Bänder verändert. Die Dispersion des N 2*s* Bandes wird erheblich vergrößert; sie beträgt 3.4 eV in GWA. Das Band erfährt insbesondere am Rand der Brillouinzone eine starke energetische Anhebung. Die Ga 3*d* Bänder werden hingegen um –2.5 eV abgesenkt. In der resultierenden Quasiteilchen-Bandstruktur beträgt die Energie der Ga 3*d* Bänder –15.7 eV. Sie liegen somit im Bereich des Minimums des N 2*s* Bandes, dessen Energie von –16.1 eV bis –12.7 eV reicht.

Ding et al. beobachten in ihrem Photoemissions-Spektrum zwei Zustände bei –17.7 eV und bei –14.2 eV [Din96]. Aus dem Intensitätsverhältnis der beiden Peaks schließen die Autoren, daß es sich bei dem tieferen der beiden Peaks um Ga 3*d* und bei dem höheren um N 2*s* handelt. Diese Zuordnung wird durch unsere Quasiteilchen-Bandstruktur bestätigt. Aufgrund unserer Quasiteilchen-Bandstruktur erwarten wir deutliche Beiträge des N 2*s* Bandes zum Photoemissionsspektrum bei den Energien am Γ-, *L*- und *X*-Punkt, also bei –16.1 eV, –13.3 eV und –12.7 eV. Unser Ergebnis für die Energien der Ga 3*d* Zustände (im Mittel –15.7 eV) liegt um 2 eV

Pseudopotentiale für Si, Ge und Sn

	ohne Semicore		mit Semicore		
	PP	Zustände	PP	Zustände	
Si	4+	$3s$, $3p$	12+	$2s$, $2p$;	$3s$, $3p$
Ge	4+	$4s$, $4p$	22+	$3s$, $3p$, $3d$;	$4s$, $4p$
Sn	4+	$5s$, $5p$	22+	$4s$, $4p$, $4d$;	$5s$, $5p$

Tabelle 13.6: *Struktur der Pseudopotentiale für Si, Ge und Sn. Wir geben an, welche Elektronen jeweils als Valenzelektronen behandelt werden.*

über dem experimentellen Befund. Wir erinnern in diesem Zusammenhang an die analogen Ergebnisse für CdS und ZnS. Wir werden das Problem der Energien stark lokalisierter Zustände in Kap. 14 detailliert diskutieren.

13.4 Elementhalbleiter

In diesem Abschnitt untersuchen wir den Einfluß der Semicore-Elektronen auf die Bandstrukturen der Elementhalbleiter Si, Ge und α-Sn. Im Hinblick auf unsere Fragestellung merken wir zunächst folgende Unterschiede zu den bisher diskutierten Verbindungshalbleitern an.

(i) Im Vergleich zu den Elementen Cd, Zn und Ga besitzen die Elemente der Gruppe IV zwei bzw. ein Valenzelektron mehr. Damit zusammenhängend, weisen die Semicore-Zustände tiefere Energien auf (LDA für Atome: E(Si $2p$) = –95.6 eV, E(Ge $3d$) = –29.6 eV, E(Sn $4d$) = –25.9 eV) als in Cd, Zn oder Ga. Die zugehörigen Wellenfunktionen sind stärker lokalisiert. Daher ist zu erwarten, daß die Semicore-Elektronen die Bandstruktur der Elementhalbleiter weniger stark beeinflussen als bei den Verbindungshalbleitern CdS, ZnS und GaN.

(ii) Bei den Verbindungshalbleitern weisen das VBM und das CBM unterschiedliche orbitale Zusammensetzungen mit verschiedenen Beiträgen vom Kation bzw. vom Anion auf. Das VBM ist eher dem Anion und das CBM eher dem Kation zuzuordnen. Die Berücksichtigung der kationischen Semicore-Zustände hat daher einen direkten Einfluß auf das Gap zwischen den beiden Zuständen. Diese räumliche Differenzierung entfällt bei den Elementhalbleitern. Wenn bei ihnen das Gap durch die Semicore-Zustände verändert wird, so nur deshalb, weil die Symmetrie der VBM-

Ausdehnung der Orbitale in Elementen der Gruppe IV

	Semicore-Radius [a.u.]	Valenz-Radius [a.u.]	$r_{\mathrm{Val}}/r_{\mathrm{Semic}}$
C	$(1s)$ 0.27	$(2s)$ 1.58	5.9
Si	$(2s)$ 0.57	$(3s)$ 2.16	3.8
Ge	$(3d)$ 0.74	$(4s)$ 2.11	2.9
Sn	$(4d)$ 1.10	$(5s)$ 2.38	2.2

Tabelle 13.7: *Radiale Ausdehnung der Semicore- und Valenz-Zustände in C, Si, Ge und Sn. Als Maß für die Ausdehnung dient der Erwartungswert des Radius:* $\int r|\psi(\mathbf{r})|^2 d^3r$. *Angegeben ist jeweils der am weitesten ausgedehnte Semicore-Zustand und der am wenigsten ausgedehnte Valenz-Zustand sowie das Verhältnis zueinander.*

und CBM-Zustände unterschiedlich ist. Bei Si und Ge ist das VBM p-artig, während das CBM hauptsächlich s-artigen Charakter besitzt. Insgesamt erwarten wir einen geringeren Einfluß der Semicore-Zustände auf die Bandstrukturen als bei den Verbindungshalbleitern.

Wir führen im Folgenden LDA- und GW-Rechnungen mit Pseudopotentialen durch, die die komplette Semicore-Schale als Valenzzustände behandeln, und vergleichen mit herkömmlichen Rechnungen ohne explizite Berücksichtigung der Semicore-Elektronen. In Tabelle 13.6 geben wir die entsprechenden Pseudopotentiale an.

Wir analysieren zuerst die atomaren Wellenfunktionen der drei Elemente. Für die Einflüsse der Semicore-Elektronen auf die Valenzbandstruktur spielen die räumlichen Ausdehnungen der Elektronen eine entscheidende Rolle; besonders das räumliche Verhältnis der Semicore- zu den Valenzelektronen ist von Bedeutung. Als repräsentative Größe untersuchen wir die mittlere räumliche Ausdehnung der Wellenfunktionen. Wir geben in Tabelle 13.7 diese Werte sowie ihr Verhältnis zueinander an. Der Vollständigheit halber schließen wir auch Kohlenstoff mit ein, obwohl wir Diamant als Kristall nicht untersuchen werden.

Es stellt sich heraus, daß bei C und Si die Valenz-Zustände wesentlich weiter ausgedehnt sind als die Semicore-Zustände. Bei Ge und Sn hingegen sind die Valenz- und Semicore-Zustände in ähnlichen Raumbereichen angeordnet. Das liegt zum einen an der generellen Vergrößerung des Atoms bei zunehmender Elektronenzahl, zum anderen an dem Auftreten von d-Elektronen in diesen Elementen. Dementsprechend erwarten wir, daß der Einfluß der Semicore-Zustände auf die Bandstruktur bei Ge

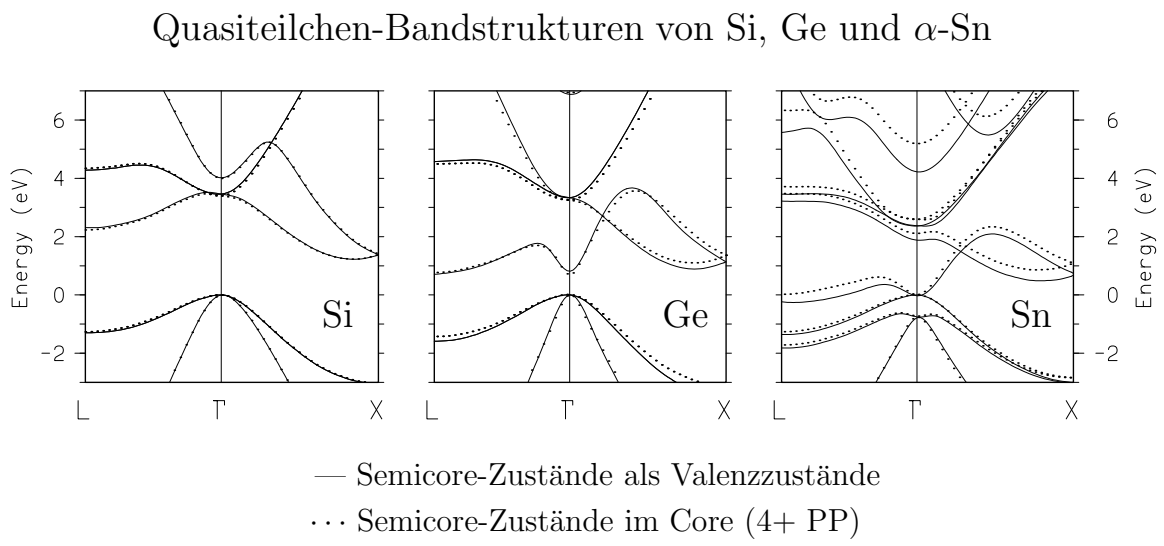


Abbildung 13.14: *GW Quasiteilchen-Bandstrukturen für Si (links), Ge (Mitte) und α-Sn (rechts). Wir vergleichen Bandstrukturen, bei deren Berechnung die Semicore-Zustände explizit berücksichtigt wurden (durchgezogene Kurven), mit Rechnungen, die mit 4+ Pseudopotentialen durchgeführt wurden (gepunktet; Semicore-Elektronen im Core). Für Sn wurde die Spin-Bahn Wechselwirkung berücksichtigt (siehe Text).*

und Sn deutlich größer ausfällt als bei Diamant oder Si.

Wir haben für Si die Bandstruktur sowohl mit dem Si^{4+} PP als auch mit dem Si^{12+} PP bestimmt. In Abb. 13.14 (links) stellen wir die zugehörigen GW-Quasiteilchen-Bandstrukturen im Bereich des fundamentalen Gaps dar. Die gepunkteten Kurven stellen die Ergebnisse ohne explizite Berücksichtigung der Semicore-Zustände dar (Si^{4+} PP), die durchgezogenen Kurven markieren die Resultate unter Mitnahme der Semicore-Zustände (Si^{12+} PP). Wie erwartet, unterscheiden sich die beiden Bandstrukturen praktisch nicht. Die Si $2s$ und $2p$ Semicore-Elektronen ändern ihre räumlichen Eigenschaften durch die Festkörper-Umgebung nicht. Außerdem sind wegen des geringen räumlichen Überlapps der Semicore-Elektronen mit den Valenz-Elektronen beide Terme der Differenz $\Delta\Sigma_{xc}$ in Gl. (13.4) für die Valenzelektronen sehr gering; somit ist auch der entprechende systematische Fehler beim Übergang von der LDA zur GWA schon bei der Si^{4+} Rechnung sehr klein. Für Diamant erwarten wir eine noch bessere Übereinstimmung der C^{4+} und C^{6+} PP Bandstrukturen. Wir verzichten daher auf die Durchführung der entsprechenden Rechnungen.

Bei Ge ergeben sich bereits geringe Unterschiede zwischen den Quasiteilchen-

Ge: Gap-Energien

[eV]	Ge^{4+}	Ge^{22+}	exp.
E_{gap}^{dir} (Γ-Γ)	<u>0.68</u>	0.81	0.89
E_{gap}^{ind} (Γ-L)	0.75	<u>0.71</u>	<u>0.744</u>

Tabelle 13.8: *Direkte (Γ-Γ) und indirekte (Γ-L) GW Quasiteilchen-Übergangsenergien für Ge. Wir geben die mit dem Ge^{4+} PP und Ge^{22+} PP berechneten Werte und experimentelle Daten an.*

Bandstrukturen der Ge^{4+} PP und der Ge^{22+} PP Rechnung (siehe mittlere Figur in Abb. 13.14). Im Bereich des Gaps differieren die beiden Bandstrukturen um bis zu 0.2 eV. Am Gap selber tritt durch die Semicore-Korrekturen eine Änderung ein (siehe Tabelle 13.8). In der herkömmlichen Ge^{4+} PP Rechnung ergibt sich eine Quasiteilchen-Bandstruktur mit einem direkten fundamentalen Gap am Γ-Punkt (0.68 eV). Am L-Punkt besitzt das unterste Leitungsband eine etwas höhere Energie von 0.75 eV. Durch die Berücksichtigung der Semicore-Zustände wird dieses Band am L-Punkt etwas abgesenkt und am Γ-Punkt angehoben. Das Minimum des Bandes befindet sich nun – ebenso wie im Experiment – am L-Punkt. Dadurch ist die Bandstruktur nunmehr indirekt halbleitend mit einem fundamentalen Gap von 0.71 eV, das sehr gut mit dem experimentellen Befund von 0.744 eV übereinstimmt. Ähnliche Veränderungen der Bandstruktur wurden von Shirley et al. durch die Berücksichtigung der Core-Polarisation erzielt [Shi92]. Rieger und Vogl [Rie95] beobachten einen Übergang von der indirekten zur direkt halbleitenden Ge Bandstruktur durch SIC-Korrekturen. Die Autoren führen dies auf die Kontraktion der Core-Orbitale durch die SIC-Korrektur zurück. In unserer Ge^{22+} PP Rechnung beobachten wir hingegen keine nennenswerte Kontraktion der Semicore-Elektronen im Vergleich zum Ge Atom.

Wir beobachten in der GW-Rechnung für Ge einen sehr ungewöhnlichen Effekt bezüglich der Topologie der Bänder in der Nähe des Gaps, wenn das Ge^{22+} PP verwendet wird. Wir stellen in Abb. 13.15 in einem kleinen Ausschnitt aus der Brillouinzone die Quasiteilchen-Bandstruktur (durchgezogene Kurven) der LDA-Bandstruktur (gestrichelt) gegenüber. Die Bandstrukturen der oberen Figur wurden mit dem Ge^{4+} PP berechnet, die der unteren Figur mit dem Ge^{22+} PP. Eine korrekte Anordnung der Bänder, die mit experimentellen Befunden verträglich ist, ist

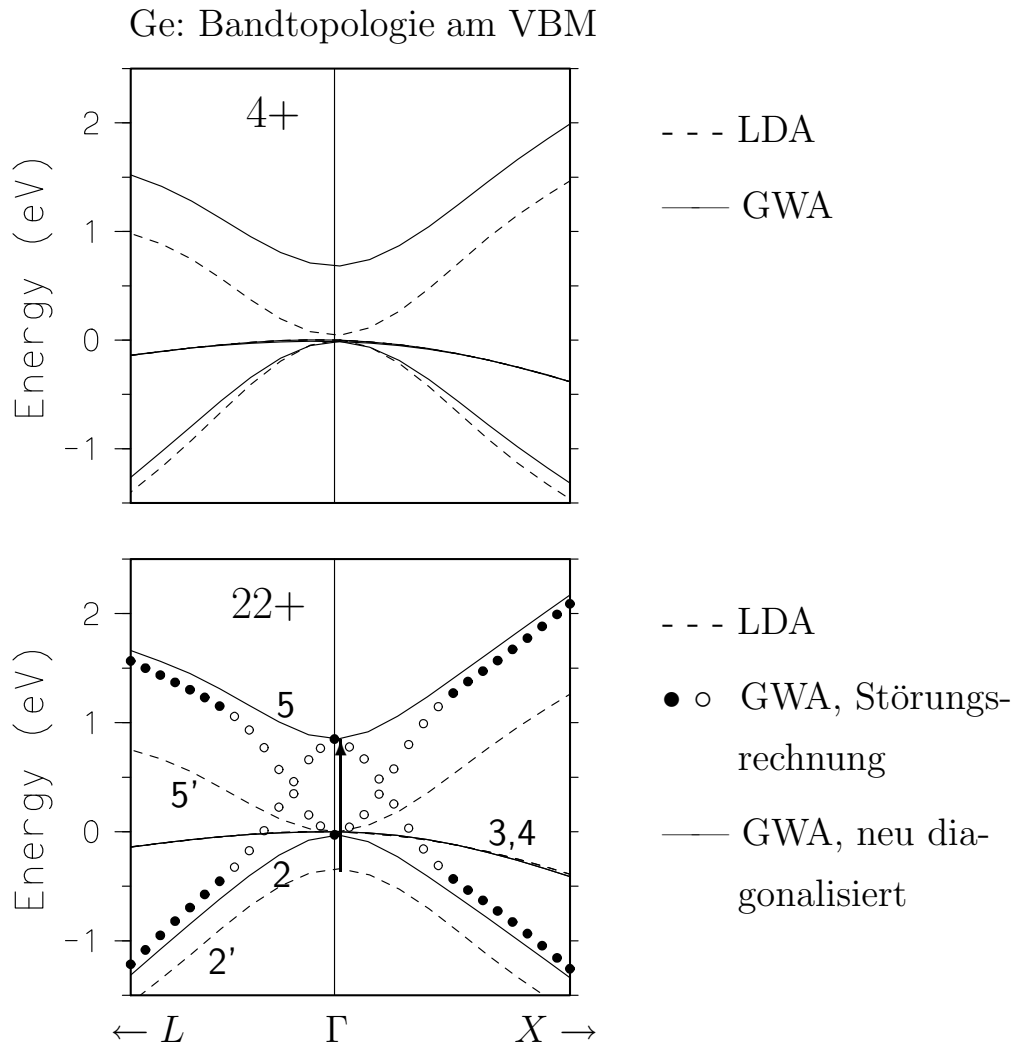


Abbildung 13.15: *Berechnete Bandstrukturenergien für Ge in der Umgebung des VBM. Oben: mit dem Ge^{4+} PP bestimmt. Unten: mit dem Ge^{22+} PP berechnet. Die gestrichelten Kurven markieren LDA-Ergebnisse. In der unteren Figur entsprechen die Punkte (•) der Standard-Vorgehensweise zur Ermittlung von GW-Quasiteilchen-Bandstrukturen mittels Störungtheorie in erster Ordnung. Dieses Verfahren ist in einigen Bereichen der Brillouinzone nicht anwendbar (offene Kreise ○). Sinnvolle Bandstruktur-Ergebnisse werden hier erst durch explizite Diagonalisierung des Hamiltonians ($H^{\mathrm{LDA}} + \Sigma(E) - V_{\mathrm{xc}}$) erzielt (siehe Text).*

durch die Ge^{4+} PP Bandstrukturen gegeben. Bei Vernachlässigung der Spin-Bahn Kopplung sind die beiden oberen Valenzbänder (schwere Löcher; *hh*) entlang *L*-Γ-*X* entartet. Am Γ-Punkt sind sie zudem mit dem dritten Valenzband (leichte Löcher; *lh*) entartet. Diese drei besetzten Zustände besitzen *p*-artige Symmetrie. Das unterste Leitungsband weist hingegen *s*-artige Symmetrie auf.

Die mit dem Ge^{22+} PP berechnete LDA-Bandstruktur zeigt in der Nähe des Γ-Punktes ein völlig anderes Verhalten. Das unbesetzte Leitungsband (5') besitzt zwar in weiten Teilen der Brillouinzone die richtige, s-artige Symmetrie, aber zum Γ-Punkt hin verändert es sich zu einem p-artigen Band. Am Γ-Punkt selbst ist es vollständig p-artig. Es hat dort die Wellenfunktion des lh-Zustands übernommen und ist energetisch mit den beiden hh-Bändern entartet. Umgekehrt wird das lh-Band (2') zum Γ-Punkt hin immer mehr s-artig. Am Γ-Punkt selbst entspricht es in seiner Symmetrie dem eigentlich unbesetzten s-artigen Leitungsbandzustand. Es besitzt eine tiefere Bandstrukturenergie als die p-Bänder.

In der GW-Rechnung, die auf dieser LDA-Rechnung basiert, stellt sich heraus, daß der Störoperator $(\Sigma(E) - V_{\rm xc})_{\rm nn'}(\mathbf{k})$ für $\mathbf{k}$-Vektoren in der Umgebung des Γ-Punktes (außer am Γ-Punkt selber) nicht diagonal ist. Er weist große Nichtdiagonal-Elemente in den LDA-Zuständen (nn')=(2′5′) bzw. (5′2′) auf. Bei derartig großen Nichtdiagonalelementen kann ein einfacher störungstheoretischer Zugang nicht angewendet werden. Statt dessen stellen wir den vollständigen Hamilton-Operator in der Basis auf, die durch die LDA-Zustände gegeben ist:

$$\begin{aligned} H^{\rm GWA}_{\rm nn'}(\mathbf{k};E) &= (H^{\rm LDA} + \Sigma(E) - V_{\rm xc})_{\rm nn'}(\mathbf{k}) \\ &= E^{\rm LDA}_{\rm n\mathbf{k}}\delta_{\rm nn'} + (\Sigma(E) - V_{\rm xc})_{\rm nn'}(\mathbf{k}) \quad . \end{aligned} \tag{13.8}$$

Für die Energieabhängigkeit des Selbstenergie-Operators verwenden wir die lineare Entwicklung (2.17) um die Energie $\bar{E} = E^{\rm LDA}_{\rm n\mathbf{k}}$. Dieser Hamiltonian ist nun zu diagonalisieren. Dabei ist seine Energie-Abhängigkeit zu beachten. Die Eigenwerte hängen also ihrerseits implizit von der Energie ab. Die endgültigen Quasiteilchen-Energien müssen die Selbstkonsistenz-Bedingung

$$H^{\rm GWA}(\mathbf{k};E^{\rm QP}_{\rm m\mathbf{k}})D_{\rm m\mathbf{k}} = E^{\rm QP}_{\rm m\mathbf{k}}D_{\rm m\mathbf{k}} \tag{13.9}$$

erfüllen. Die zugehörigen Quasiteilchen-Amplituden ergeben sich als Linearkombinationen der LDA-Wellenfunktionen:

$$\psi^{\rm QP}_{\rm m\mathbf{k}} = \sum_{\rm n'} D^{\rm n'}_{\rm m\mathbf{k}}\psi^{\rm LDA}_{\rm n'\mathbf{k}} \quad . \tag{13.10}$$

Da Nichtdiagonal-Elemente nur zwischen 2′ und 5′ auftreten und der Hamiltonian ansonsten diagonal ist, können wir uns bei der Diagonalisierung auf den von 2′ und 5′ aufgespannten Unterraum beschränken. Alle anderen Quasiteilchen-Energien lassen

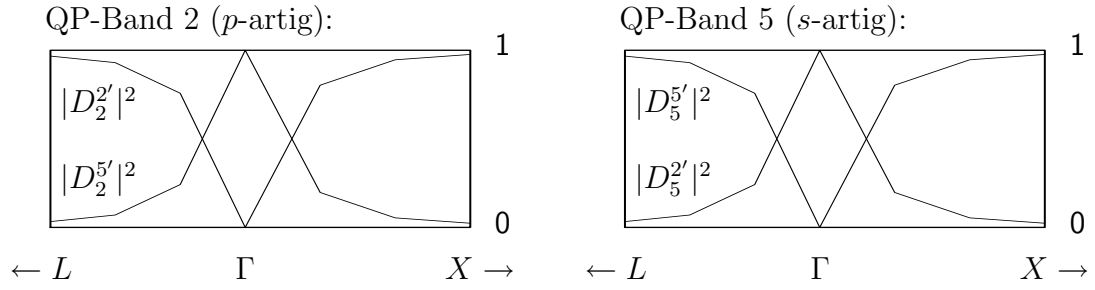


Abbildung 13.16: *Eigenvektor-Koeffizienten $D^{\mathrm{n}'}_{\mathrm{m}\mathbf{k}}$ der expliziten Diagonalisierung des GWA-Hamiltonians für Ge (siehe Text).*

sich durch Störungsrechnung erster Ordnung ermitteln; ihre Amplituden sind durch die LDA-Wellenfunktionen gegeben. Die auf diese Weise bestimmten Quasiteilchen-Bänder sind in Abb. 13.15 (unten) als durchgezogene Kurven dargestellt. Die Abbildung suggeriert, daß das LDA-Band $2'$ starr zu höheren Energien verschoben worden sei ($\rightarrow 2$), ebenso das LDA-Band $5'$ ($\rightarrow 5$). Das ist nicht der Fall. Am linken und rechten Kastenrand sind die Nichtdiagonalelemente klein, und es tritt nur eine starre Energie-Verschiebung auf; die LDA-Wellenfunktionen bleiben dabei erhalten. In der Umgebung des Γ-Punkts ist jedoch die scheinbare Zuordnung $2' \rightarrow 2$ und $5' \rightarrow 5$ gar nicht möglich. In diesem Bereich sind die beiden LDA-Zustände $2'$ und $5'$ keine Eigenzustände des GW-Hamiltonians, sondern sie werden linear zu völlig neuen Quasiteilchen-Zuständen kombiniert. Wir stellen in Abb. 13.16 den Verlauf der Koeffizienten $D^{\mathrm{n}'}_{\mathrm{m}\mathbf{k}}$ in diesem Bereich dar. Lediglich am Γ-Punkt selbst ist wieder eine Zuordnung $2' \rightarrow 5$ und $5' \rightarrow 2$ vorhanden (in Abb.13.15 als Pfeil angedeutet).

Die Ergebnisse der einfachen GW Störungsrechnung erster Ordnung sind in Abb. 13.15 (unten) als Punkte und offene Kreise dargestellt. Der Vergleich mit den konsistent bestimmten GW Quasiteilchen-Energien zeigt, daß am linken und rechten Kastenrand diese Störungsrechnung ohne weiteres durchgeführt werden kann und sinnvolle Resultate für die Bandstruktur ergibt. In diesem Bereich liefert ja auch die vollständige Diagonalisierung Quasiteilchen-Amplituden, die mit den LDA-Wellenfunktionen sehr gut übereinstimmen (vgl. Abb. 13.16). In der Umgebung des Γ-Punktes führt die Störungsrechnung jedoch auf Ergebnisse, die mit denen der Diagonalisierung nicht übereinstimmen. Hier ist die einfache Störungsrechnung nicht anwendbar. Wir haben diese Daten daher als offene Kreise gekennzeichnet.

Die hier vorgestellten Einzelheiten der Quasiteilchen-Korrekturen sind letztlich

nicht von entscheidender Bedeutung. Von zentraler Wichtigkeit ist hingegen, *daß die GWA imstande ist, topologische Fehler der LDA-Bandstruktur vollständig zu korrigieren.* Es war in keiner Weise zu erwarten, daß der Selbstenergie-Operator – der ja letztlich auf LDA-Ergebnissen beruht – die LDA-Zustände, obwohl sie in der Umgebung des Γ-Punktes keine echten Eigenzustände des physikalischen Systems sind, zu den richtigen Quasiteilchen zusammenzusetzen vermag. Die Tatsache, daß die Quasiteilchen-Bänder die richtige Reihenfolge besitzen und die Symmetrie der zugehörigen Zustände der Realität entspricht, obwohl das in der zugrunde liegenden LDA-Bandstruktur nicht der Fall war, verleiht dem gesamten GW-Zugang zur Berechnung von Bandstrukturen eine erhebliche Zuverlässigkeit.

Wir erinnern in diesem Zusammenhang daran, daß auch bei GaN Struktur und Zusammensetzung der Zustände durch die Quasiteilchen-Korrekturen stark verändert worden waren (siehe letzten Abschnitt). Eine selbstkonsistente Bestimmung des Selbstenergie-Operators war auf der Grundlage der reinen LDA nicht möglich gewesen. Dieses Problem trifft prinzipiell auch auf den vorliegenden Fall der Ge Bandstruktur zu. Hier betrifft die Umstrukturierung der Zustände jedoch nur einen sehr kleinen Bereich nahe des Γ-Punktes, nicht aber die gesamte Brillouinzone. Diesem kleinen Bereich kommt insgesamt nur sehr wenig Bedeutung zu. Im überwiegenden Teil der Brillouinzone ändern sich die Zustände durch die Quasiteilchen-Korrekturen nicht. Daher läßt sich für Ge der Selbstenergie-Operator schon auf der Basis der LDA nahezu selbstkonsistent bestimmen.

Wir merken an, daß die Bandstruktur-Energie der Ge $3d$ Zustände weit unterhalb des untersten Valenzbandes liegt. Ihre Quasiteilchen-Energie beträgt im Mittel – 27.7 eV. Eine direkte Kopplung der $3d$ Zustände an die Valenzbänder, wie etwa in CdS, ZnS und GaN, beobachten wir bei Ge nicht.

Zum Abschluß dieser systematischen Untersuchung der Gruppe IV Halbleiter betrachten wir Zinn, das in seiner metastabilen Modifikation α-Sn (graues Zinn) in der Diamantstruktur vorliegt. Von Diamant bis Zinn nimmt die elektronische Bandlücke der Kristalle immer weiter ab. Bei α-Sn ist die Bandlücke exakt Null, und das Material ist ein Halbmetall [Che76, Bru93]. In der LDA ergibt sich zunächst eine *metallische* Bandstruktur für α-Sn, in der das unterste Leitungsband am L-Punkt eine niedrigere Energie aufweist als das Valenzband-Maximum am Γ-Punkt. Wir wollen hier untersuchen, ob dies in einer Quasiteilchen-Rechnung korrigiert wird.

Die von C bis Sn zunehmende Kernladungszahl führt dazu, daß die Spin-Bahn Wechselwirkung als relativistischer Effekt größer wird. Bereits bei Ge führt sie zu

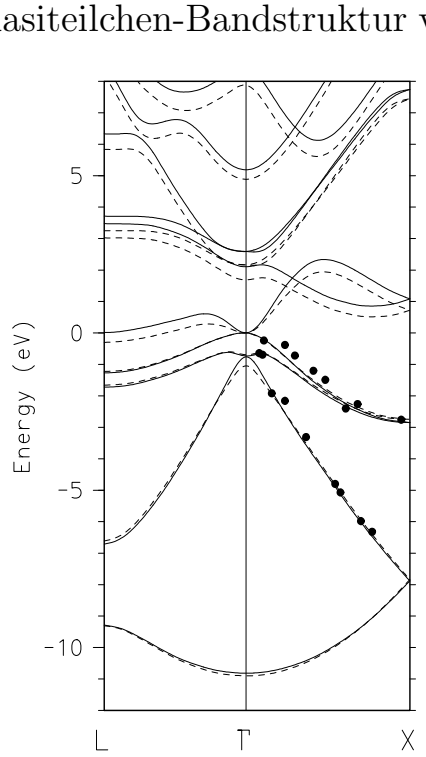


Abbildung 13.17: *GW Quasiteilchen-Bandstrukturen von α-Sn im Vergleich zur LDA-Bandstruktur (gestrichelt). Es wurde das Sn^{4+} PP verwendet. In der Rechnung wurde die Spin-Bahn Wechselwirkung explizit berücksichtigt [Wür95, Wür96]. Die experimentellen Daten wurden von Höchst und Hernández-Calderón mittels ARUPS gewonnen [Höc83].*

Aufspaltungen der Bandstruktur-Energien von bis zu 0.3 eV [Hyb86]. Bei Zinn fällt dieser Einfluß noch stärker aus. Für eine genaue Bestimmung der Topologie der Bänder am VBM ist eine Berücksichtigung der Spin-Bahn Wechselwirkung unerläßlich. Im Rahmen unseres Pseudopotential-Zugangs kann das in guter Näherung durch die Mitnahme entsprechender Pseudopotentiale für den Spin-Bahn Kopplungsoperator geschehen [Bac82, Wür95]. Die Theorie der Spin-Bahn Wechselwirkung soll hier nicht weiter vertieft werden. Zu einer ausführlichen Darstellung verweisen wir auf die Arbeit von Würde [Wür96]. Zur Berücksichtigung der Spin-Bahn Kopplung muß der gesamte Formalismus um die Spinvariablen erweitert werden, die wir bisher nur im Rahmen der simplen Spinentartung behandelt hatten. Diese Erweiterung betrifft sowohl die LDA als auch die GWA. Sie führt zu einer Verdoppelung der Dimension des Hamiltonians.

Wir haben zunächst Rechnungen mit einem Sn^{4+} PP durchgeführt. Dabei wurde in der LDA das korrigierende Pseudopotential zur Spin-Bahn Wechselwirkung berücksichtigt. Darauf aufbauend, wurden Quasiteilchen-Korrekturen im Rahmen der GWA bestimmt. Wir stellen in Abb. 13.17 die resultierende Quasiteilchen-Bandstruktur im direkten Vergleich zur LDA-Bandstruktur (gestrichelt) und zu experimentellen ARUPS-Daten von Höchst und Hernández-Calderón [Höc83] vor. Die von uns berechneten Energien stimmen hervorragend mit den gemessenen Daten überein. Die Bandstruktur gleicht weitgehend der für Si oder Ge. Zwei Besonderheiten sind hervorzuheben.
(i) Das unterste Leitungsband weist in der LDA-Bandstruktur entlang der $L\Gamma$ Linie Energien unterhalb des VBM auf. Am L-Punkt beträgt seine LDA-Bandstrukturenergie –0.30 eV. Die Bandstruktur ist somit metallisch. Im Experiment wird hingegen eine Energie von 0.11 eV oberhalb des VBM gemessen. In der GW-Rechnung ermitteln wir für dieses Band eine Quasiteilchen-Korrektur von +0.31 eV am L-Punkt und somit eine Bandstrukturenergie von 0.01 eV in guter Übereinstimmung mit dem experimentellen Wert. Im Rahmen dieser GW-Rechnung ergibt sich für α-Sn eine halbmetallische Bandstruktur.
(ii) Das oberste Valenzband und das unterste Leitungsband nähern sich in ihrer Energie zum Γ-Punkt hin an und sind am Γ-Punkt entartet. Die zugehörigen Wellenfunktionen besitzen p-artige Symmetrie. Die exakte Entartung wird durch die Symmetrie des Kristalls verursacht. Es handelt sich dabei um den gleichen Effekt, den wir bei Ge in der LDA-Rechnung mit dem Ge^{22+} PP beobachtet und dort als Artefakt der Rechnung identifiziert hatten. Bei α-Sn hingegen entspricht diese Entartung der Realität; sie macht gerade den halbmetallischen Charakter des Materials aus. Das unterste Leitungsband erfährt durch die Quasiteilchen-Korrekturen in weiten Bereichen der Brillouinzone eine Anhebung zu höheren Energien. Am Γ-Punkt hingegen beträgt seine Korrektur Null, und die exakte Entartung mit dem obersten Valenzband an diesem Punkt bleibt auch in der GW Quasiteilchen-Bandstruktur bestehen.

Wir haben auch für α-Sn die Auswirkungen der Semicore-Elektronen auf die Quasiteilchen-Bandstruktur untersucht, die sich ergeben, wenn die gesamte Semicore-Schale als Valenzschale behandelt wird. Um den numerischen Aufwand zu begrenzen, haben wir die entsprechenden Korrekturen zwischen Sn^{4+} und Sn^{22+} in einer Rechnung ohne Spin-Bahn Wechselwirkung bestimmt. Anschließend wurden diese Semicore-Korrekturen auf die Sn^{4+} PP Quasiteilchen-Bandstruktur mit Spin-

α-Sn: Übergangs-Energien

	Sn^{4+}	Sn^{22+}	exp.
an L	4.7	4.6	4.1–4.2
an Γ	2.1	1.9	2.0–2.4
an X	3.9	3.7	3.6–3.7

Tabelle 13.9: *Berechnete Quasiteilchen-Übergangsenergien für α-Sn am L-, Γ- und X-Punkt. Wir beschränken uns auf die jeweils niedrigsten direkten Übergänge. Die experimentellen Daten sind aus [Lan82] entnommen.*

Bahn Kopplung angewendet (siehe rechte Figur in Abb. 13.14). Die Korrekturen durch den Übergang vom Sn^{4+} PP zum Sn^{22+} PP sind deutlich größer als bei Si oder Ge. Es kommt bei fast allen Zuständen zu einer Absenkung der Quasiteilchen-Energien gegenüber der Sn^{4+} GW Rechnung. Bei den Leitungsbändern fällt sie deutlich größer aus als bei den Valenzbändern.

Ein detaillierter Vergleich der berechneten mit experimentellen Daten gestaltet sich wegen der vielfältigen Spin-Bahn Aufspaltungen als schwierig. Wir diskutieren hier lediglich die niedrigsten direkten Quasiteilchen-Übergangsenergien und vergleichen mit vorliegenden experimentellen Werten (siehe Tabelle 13.9). Wir finden generell eine gute Übereinstimmung der berechneten mit den gemessenen Daten. Lediglich am L-Punkt überschätzen unsere Rechnungen die experimentell bestimmte Übergangsenergie. Generell stimmen die Übergangsenergien der Sn^{22+} PP Rechnung etwas besser mit den gemessenen Daten überein. Allerdings liegt in der Sn^{22+} PP Quasiteilchen-Bandstruktur das unterste Leitungsband am L-Punkt um 0.22 eV unter dem VBM, so daß sich eine leicht metallische Bandstruktur ergibt.

Kapitel 14

Semicore- und Core-Niveaus

Bisher standen die Einflüsse der Semicore-Elektronen auf die Valenz- und Leitungsbandstruktur im Vordergrund unserer Überlegungen. Über die Energien der Semicore-Elektronen selbst haben wir bisher noch nicht viel ausgesagt (abgesehen von den Cd $4d$ Zuständen). Da wir die Semicore-Elektronen als Valenzelektronen behandeln, erhalten wir aus unserem Zugang für sie in gleicher Weise Wellenfunktionen und Bandstruktur-Energien wie für die herkömmlichen Valenzelektronen. Diesen Energien wollen wir uns nun zuwenden. Dabei stehen neben quantitativen Ergebnissen für einzelne physikalische Systeme vor allem grundlegende Erkenntnisse über die Eigenschaften und die Aussagekraft der verwendeten Methoden im Blickpunkt unseres Interesses.

Wir werden im folgenden immer die Bandstruktur-Energien der betrachteten Zustände angeben und diese auf das Volumen-Valenzbandmaximum (VBM) beziehen. Häufig werden in der Literatur auch die Bindungsenergien der Zustände angegeben. Wenn diese auf das VBM bezogen werden, so stimmen sie mit unserer Bandstrukturenergie bis auf das Vorzeichen überein. Bei einer Angabe der Bindungsenergien bezüglich des Vakuumniveaus kommt als zusätzlicher Unterschied die Ionisierungsenergie des Materials hinzu.

Die Eigenschaften der ca. 10^3 verschiedenen Semicore- und Core-Zustände der chemischen Elemente sind einander ähnlich. Wir untersuchen hier exemplarisch drei solche Zustände: Si $2p$ in Si, Ge $3d$ in Ge und Cd $4d$ in kubischem CdS. Wir diskutieren vier verschiedene theoretische Zugänge zur Berechnung der Bandstruktur-Energie von Semicore- und Core-Elektronen. Zunächst betrachten wir die LDA (1) und berechnen mit der bisher verwendeten Standard-GWA (2) Quasiteilchen-Korrekturen dazu. Danach untersuchen wir Modifikationen der Green-Funktion G,

die auf eine selbstkonsistente Behandlung der GW-Näherung hinarbeiten. Wir werden dabei zunächst das in die Green-Funktion eingehende Einteilchen-Energiespektrum selbstkonsistent gestalten (3a) und schließlich noch zusätzlich Plasmon-induzierte Satelliten-Strukturen in der Green-Funktion berücksichtigen (3a+b).

Wir geben in diesem Kapitel nur die mittleren Energien der Zustände an. Die Kristallfeld-Aufspaltung durch die im Festkörper gestörte Isotropie, die sich in unseren Rechnungen automatisch ergibt, bleibt bei der Angabe der Ergebnisse unberücksichtigt. Die Spin-Bahn Aufspaltung wird bei unseren Rechnungen vernachlässigt. Sie ist für unsere Fragestellung von untergeordneter Bedeutung und läßt sich bei Bedarf im Rahmen einer Störungsrechnung nachträglich berücksichtigen.

14.1 LDA und Standard-GWA

(1) Im Rahmen der LDA ergeben sich generell zu hohe Einteilchen-Energien für Semicore- und Core-Zustände (zu geringe Bindungsenergie). Um das zu verstehen, betrachten wir zunächst die Energie, die mit dem Entfernen eines Elektrons (n, l) aus einem isolierten Atom verbunden ist. Innerhalb der LDA lassen sich hierfür zwei verschiedene Werte angeben. Ein konzeptionell richtiges Ergebnis für die Ablösearbeit des Elektrons (n, l) erhält man aus dem Vergleich der totalen Energien des entsprechenden $(N–1)$-Elektronen Ions und des neutralen N-Elektronen Atoms. Das LDA-Einteilchen-Niveau $E_{n,l}$ im Atom ergibt hingegen ein anderes Resultat. Im Rahmen der Kohn-Sham Gleichungen betrachtet man *infinitesimal kleine* Besetzungsänderungen und berechnet die damit verbundene infinitesimale Energieänderung. Der Einteilchen-Energieeigenwert entspricht der linearen Extrapolation dieser Energieänderung auf die endliche Besetzungsänderung ± 1. Da aber das Dichtefunktional nicht linear mit der Besetzung zusammenhängt, tritt ein systematischer Unterschied zur Gesamtänderung der totalen Energie auf. So beträgt zum Beispiel der Energie-Eigenwert des Si $2p$ Zustands im neutralen Si Atom –95.61 eV, die Differenz der totalen Energien von Si und Si^+ bei Entfernen eines $2p$ Elektrons hingegen –110.10 eV. Man beachte, daß sich der Energieeigenwert selbst bei der Veränderung der Besetzung ändert. Im Si^+ Ion mit nur noch 5 Si $2p$ Elektronen beträgt der Eigenwert –125.36 eV.

Die generelle Unterschätzung der Bindungsenergie durch den Energieeigenwert betrifft auch die Valenzelektronen des Atoms. Der Fehler liegt hier in der Größenordnung von 5 eV. Vor diesem Hintergrund ist es auf den ersten Blick erstaunlich,

daß die LDA-Bandstrukturen von Oberflächensystemen dennoch einigermaßen zuverlässige Aussagen über die Ionisierungsenergie von Elektronen aus Festkörpern mit Ungenauigkeiten von meistens weniger als 1 eV erlauben (vgl. Kap. 10). Der Grund liegt darin, daß die Valenzelektronen im Festkörper zu weit ausgedehnten, delokalisierten Zuständen hybridisieren. Wird ein Elektron aus solch einem Zustand aus dem System entfernt, so ist die damit verbundene Dichteänderung an jedem einzelnen Atom sehr gering. Dieser Prozeß entspricht also weitgehend der Annahme einer lokal infinitesimal kleinen Dichteänderung, die den Kohn-Sham Gleichungen zugrunde liegt. Daher stellen die Kohn-Sham Energien für die Valenz- und Leitungsbandstruktur ein in dieser Hinsicht konsistentes Ergebnis dar – abgesehen von den Quasiteilchen-Korrekturen, die dem Unterschied zwischen Grundzustands-Veränderungen und teilchenartigen Anregungen Rechnung tragen. Auf die Core-Elektronen trifft die Aussage über die Infinitesimalität der Dichteänderung nicht zu. Anregungen solcher Zustände lassen sich eher als an einem Atom lokalisiert beschreiben. Damit unterliegt der zugehörige LDA-Energieeigenwert auch im Festkörper dem oben erwähnten systematischen Fehler bezüglich der endlichen Dichteänderung und unterschätzt die Bindungsenergie des Zustandes.

Wir geben in Tabelle 14.1 die LDA-Eigenwerte der drei untersuchten Zustände in den Kristallen Si, Ge und CdS an (Spalte 1). Sie liegen offensichtlich weit über den experimentellen Daten, d.h. ihre LDA-Bindungsenergie ist zu klein.

(2) Die GW-Näherung liefert Resultate für Quasiteilchen-Energien, die per Konstruktion mit dem betrachteten Abtrennungsprozeß eines Elektrons konsistent sind. Der Green-Funktions Formalismus wurde genau dafür entwickelt. Er ermöglicht die konzeptionell richtige Bestimmung der Selbstenergie von Core-Zuständen. Während der Austausch-Anteil im wesentlichen durch die Austauschenergie im Atom gegeben ist, unterscheidet sich der Korrelationsanteil deutlich vom Atom. Der Unterschied liegt darin begründet, daß die Polarisierbarkeit und damit die Abschirmung der Wechselwirkung im Festkörper völlig anders gestaltet ist als im Atom. Erste Berechnungen der Selbstenergie von Core-Zuständen wurden von Bechstedt et al. durchgeführt [Bec81].

Unsere mit der Standard-GWA erzielten Resultate für die Semicore- und Core-Niveaus zeigen gegenüber der LDA eine deutliche Energieabsenkung (siehe Spalte 2 in Tabelle 14.1). Die Abweichungen der LDA zu den experimentellen Energiedaten werden in etwa halbiert. Dennoch liegen auch diese Standard GW-Ergebnisse für die Semicore- und Core-Energien deutlich über den experimentellen Werten. Core-

Quasiteilchen-Energien von Si $2p$, Ge $3d$ und Cd $4d$

[eV]	LDA	GWA			Exp.
$G(\omega)$		E^{LDA}	E^{QP}	E^{QP} + Satell.	
	(1)	(2)	(3a)	(3a+b)	
Si $2p$	–89.4	–94.8	–97.9	–100.4	–100.0[a], –99.0[b]
Ge $3d$	–24.6	–27.7	–28.2	–30.0	–29.7[a], –29.5[b]
Cd $4d$	–7.2	–8.1	–8.2	–9.1	–9.5[c], –9.2[d]

[a]Ref. [Mön93] [b]Ref. [Lan89] [c]Ref. [Ley74, Sto83] [d]Ref. [Sta96]

Tabelle 14.1: *Quasiteilchen-Energien der Core-Zustände Si 2p in Si, Ge 3d in Ge und Cd 4d in kubischem CdS. Die Ergebnisse von vier Methoden werden miteinander verglichen: (1) LDA-Eigenwerte, (2) Standard-GWA, (3a, 3a+b) Erweiterungen der GWA (siehe Text).*

Niveaus wurden auch von Aryasetiawan und Gunnarsson im Rahmen der GWA bestimmt [Ary96]. Auch die Ergebnisse dieser Autoren liegen über den experimentellen Daten.

14.2 Erweiterungen der GWA

Eine Ursache für die Diskrepanz zwischen den Core-Niveau Ergebnissen der Standard-GWA und den gemessenen Werten liegt darin, daß in der Standard-GWA nur eine einzige Iteration bezüglich des Selbstenergie-Operators durchgeführt wird:

$$\Sigma^0 = G^0 W^0 \tag{14.1}$$

Der Index „0“ deutet an, daß in G und W nur die Ergebnisse der zugrunde liegenden LDA Rechnung eingehen, nicht aber die korrigierten Daten, die aus der GWA folgen. Wir werden nun an der Green-Funktion G, die in diesen Operator linear eingeht, Modifikationen in Richtung auf einen selbstkonsistenten Selbstenergie-Operator vornehmen:

$$\Sigma = G W^0 \tag{14.2}$$

In die abgeschirmte Wechselwirkung W gehen hauptsächlich solche Zustände ein, die eine starke Polarisierbarkeit aufweisen, also die Valenz- und Leitungsbandzustände. Die Polarisierbarkeit dieser Zustände wird durch die LDA hinreichend gut beschrieben. Die Core-Zustände, deren Energien wir untersuchen, haben nur einen geringen Einfluß auf die abgeschirmte Wechselwirkung (vgl. Abschnitt 13.2). Wenn in der Green-Funktion Modifikationen an den Energien dieser Core-Zustände vorgenommen werden, so wird die abgeschirmte Wechselwirkung davon nur unwesentlich berührt. Wir können sie also weiterhin in guter Näherung aus LDA-Ergebnissen berechnen: $W = W^0$. Wir wollen im folgenden zwei der wichtigsten Aspekte der Selbstkonsistenz beleuchten. Dabei konzentrieren wir uns auf die Bedeutung des eingehenden Einteilchen-Spektrums und auf die Energieabhängigkeit der Selbstenergie. Beide Aspekte betreffen den Korrelationsanteil der Selbstenergie. Die Austausch-Energie wird davon nicht berührt und ist mit der Austausch-Energie der Standard-GWA (2) identisch.

(3a) Für die Quasiteilchen-Korrektur eines Zustandes $|\mathrm{m}\mathbf{k}\rangle$ muß grundsätzlich der selbstkonsistente Selbstenergie-Operator an der resultierenden Quasiteilchen-Energie ausgewertet werden:

$$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = \langle \Sigma(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}} \quad . \tag{14.3}$$

[Unter Selbstkonsistenz verstehen wir in diesem Absatz zunächst nur, daß die in den Selbstenergie-Operator eingehenden Einteilchen-Energien mit den resultierenden Quasiteilchen-Energien übereinstimmen sollen.] Es stellt sich heraus, daß die Quasiteilchen-Korrekturen delokalisierter Valenz- und Leitungsbänder in sehr guter Näherung auch mit dem nicht selbstkonsistenten Selbstenergie-Operator Σ^0 bestimmt werden können, in den lediglich die nicht korrigierten LDA-Energien eingehen:

$$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = \langle \Sigma^0(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}} \quad . \tag{14.4}$$

Wir haben für Si beide Vorgehensweisen verglichen [Roh93a]. Die Unsicherheiten durch die vereinfachte Vorgehensweise (14.4) belaufen sich auf maximal 0.15 eV; im Gap-Bereich (von –5 eV bis +5 eV) betragen sie maximal 0.05 eV. Darüber hinaus läßt sich die Auswertung von (14.4) an der Quasiteilchen-Energie $E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}$ in hervorragender Weise ersetzen durch die Auswertung an der LDA-Energie $E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}$ und nachfolgende Multiplikation mit der Renormierungskonstante $Z_{\mathrm{m}\mathbf{k}}$:

$$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = Z_{\mathrm{m}\mathbf{k}} \langle \Sigma^0(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}} \quad . \tag{14.5}$$

		delokal. Val./Leit.	Core-Zustand
	exakt:		
(14.3)	$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = \langle \Sigma(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}}$	•	•
	kann ersetzt werden durch:		
(14.4)	$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = \langle \Sigma^0(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}}$	•	—
(14.5)	$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = Z_{\mathrm{m}\mathbf{k}} \langle \Sigma^0(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}}$	•	—
(14.7)	$\Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} = \langle \Sigma^0(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}}$	—	•

Tabelle 14.2: *Anwendbarkeit vereinfachter Rechenvorschriften zur Berechnung von Quasiteilchen-Korrekturen im Falle delokalisierter Valenz- und Leitungszustände sowie für lokalisierte Core-Zustände (siehe Text).*

Dies ist genau die Rechenvorschrift, die wir bisher verwendet hatten. Die verschiedenen Vorgehensweisen sind in Tabelle 14.2 zusammengefaßt. Die gute Übereinstimmung zwischen (14.3) und (14.4) für Valenz- und Leitungszustände beruht unter anderem darauf, daß die Quasiteilchen-Korrekturen der Zustände relativ gering sind. Außerdem setzt sich ihre Selbstenergie aus Beiträgen von Valenz- sowie Leitungszuständen zusammen, die ja gerade entgegengesetzte Korrekturen erfahren. Die Korrektur des Spektrums beeinflußt daher die Selbstenergien dieser Zustände kaum. Im Falle eines Core-Elektrons treffen beide Voraussetzungen nicht mehr zu. Ein Core-Zustand erfährt eine erhebliche Quasiteilchen-Korrektur, und seine Selbstenergie kommt hauptsächlich durch Kopplung des Zustands mit sich selbst zustande. Daher ist die Vereinfachung (14.3)→(14.4) für die Quasiteilchen-Korrektur eines Core-Zustands nicht zulässig. Aber auch in diesem Fall existiert eine vereinfachte Rechenvorschrift. Diese stellen wir im Folgenden kurz vor.

Wir betrachten einen Zustand $|\mathrm{m}\mathbf{k}\rangle$, der eine Quasiteilchen-Korrektur Δ erfährt. Wir nehmen an, daß die Korrelationsenergie dieses Zustands hauptsächlich durch Kopplung mit solchen Zuständen $|\mathrm{n}\mathbf{q}\rangle$ herrührt, die die gleiche Quasiteilchen-Korrektur Δ erfahren. Diese Annahme ist bei den Core-Elektronen in guter Näherung erfüllt, denn die Korrelationsenergie eines Core-Zustands kommt hauptsächlich durch die Kopplung des Zustands mit sich selbst zustande (vgl. Tabelle 13.4 für die Cd 4*d* Zustände in CdS). Wir erhalten folgendes Resultat für die Selbstenergie des

Core-Zustands:

$$\begin{aligned}
\Sigma_{\mathrm{m}\mathbf{k}}(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) &= \left\langle \frac{i}{2\pi}\int e^{-i\omega 0^+} G(\mathbf{r},\mathbf{r}', E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}-\omega) W^0(\mathbf{r},\mathbf{r}',\omega) d\omega \right\rangle_{\mathrm{m}\mathbf{k}} \\
&= \left\langle \sum_{n,\mathbf{q}} \frac{i}{2\pi}\int \frac{e^{-i\omega 0^+}\psi_{\mathrm{n}\mathbf{q}}(\mathbf{r})\psi^*_{\mathrm{n}\mathbf{q}}(\mathbf{r}')}{E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} - E^{\mathrm{QP}}_{\mathrm{n}\mathbf{q}} - \omega \pm i0^+} W^0(\mathbf{r},\mathbf{r}',\omega) d\omega \right\rangle_{\mathrm{m}\mathbf{k}} \\
&= \left\langle \sum_{n,\mathbf{q}} \frac{i}{2\pi}\int \frac{e^{-i\omega 0^+}\psi_{\mathrm{n}\mathbf{q}}(\mathbf{r})\psi^*_{\mathrm{n}\mathbf{q}}(\mathbf{r}')}{E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} + \Delta - (E^{\mathrm{LDA}}_{\mathrm{n}\mathbf{q}} + \Delta) - \omega \pm i0^+} W^0(\mathbf{r},\mathbf{r}',\omega) d\omega \right\rangle_{\mathrm{m}\mathbf{k}} \\
&= \left\langle \sum_{n,\mathbf{q}} \frac{i}{2\pi}\int \frac{e^{-i\omega 0^+}\psi_{\mathrm{n}\mathbf{q}}(\mathbf{r})\psi^*_{\mathrm{n}\mathbf{q}}(\mathbf{r}')}{E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}} - E^{\mathrm{LDA}}_{\mathrm{n}\mathbf{q}} - \omega \pm i0^+} W^0(\mathbf{r},\mathbf{r}',\omega) d\omega \right\rangle_{\mathrm{m}\mathbf{k}} \\
&= \left\langle \frac{i}{2\pi}\int e^{-i\omega 0^+} G^0(\mathbf{r},\mathbf{r}', E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}-\omega) W^0(\mathbf{r},\mathbf{r}',\omega) d\omega \right\rangle_{\mathrm{m}\mathbf{k}} \\
&= \Sigma^0_{\mathrm{m}\mathbf{k}}(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}) \quad , \qquad (14.6)
\end{aligned}$$

$$\begin{aligned}
\text{also} \quad \Delta^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}} &= \langle \Sigma(E^{\mathrm{QP}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}} \\
&= \langle \Sigma^0(E^{\mathrm{LDA}}_{\mathrm{m}\mathbf{k}}) - V_{\mathrm{xc}} \rangle_{\mathrm{m}\mathbf{k}} \quad . \qquad (14.7)
\end{aligned}$$

Wir gelangen zu dem überraschenden Resultat, daß die Selbstenergie in nullter Näherung, ausgewertet bei der LDA-Energie, *ohne Multiplikation mit der Renormierungskonstanten* bereits die Quasiteilchen-Korrektur liefert, die in einem aufwendigen iterativen Verfahren letztlich bei Selbstkonsistenz erzielt wird. Ein Vergleich der damit erzielten Ergebnisse für das Si $2p$ Niveau in Si sowie für Ge $3d$ in Ge mit denen der exakten, selbstkonsistenten Vorgehensweise (14.3) ergibt, daß diese vereinfachte Rechenvorschrift (14.7) für Core-Energien mit Unsicherheiten von weniger als 0.1 eV behaftet ist.

Die Korrektur des Spektrums (=Weglassen der Renormierungskonstante) führt zu einer erheblichen Vergrößerung der Quasiteilchen-Korrektur von Core-Zuständen und zu einer zusätzlichen Absenkung ihrer Quasiteilchen-Energie gegenüber der Standard-GWA (2). Sie beträgt –3.1 eV für den Zustand Si $2p$, –0.5 eV für Ge $3d$ sowie –0.1 eV für Cd $4d$. Die resultierenden Quasiteilchen-Energien für die drei Zustände sind in Tabelle 14.1 aufgeführt (Spalte 3).

Schon durch die Standard-GWA (2) erfuhr der Core-Zustand Si $2p$ eine weit stärkere Energieabsenkung als etwa der Semicore-Zustand Cd $4d$. Durch die Kor-

rektur des Spektrums wird die Standard-GW Quasiteilchen-Korrektur (2) um den Faktor Z^{-1} vergrößert. Daher fällt die zusätzliche Energieabsenkung für den Zustand Si $2p$ wesentlich größer aus als für den Zustand Cd $4d$. Die Korrektur des Spektrums bzw. Nichtbeachtung der Renormierungskonstante führt zu einer besseren Übereinstimmung der berechneten Bandstruktur-Energien (3a) mit den experimentellen Daten als bei der Standard-GWA (2). Unsere Ergebnisse liegen aber nach wie vor im Vergleich zum Experiment bei zu hohen Energien.

(3a+b) Bei den bisherigen Überlegungen (3a) hatten wir den Elektronen eine Green-Funktion zugeordnet, die einem System effektiv unabhängiger Teilchen entspricht:

$$G(\mathbf{x},\mathbf{x}',E) = \sum_{\mathbf{nq}} \psi_{\mathbf{nq}}(\mathbf{x})\psi^*_{\mathbf{nq}}(\mathbf{x}') \cdot \frac{1}{E - E_{\mathbf{nq}} \pm i0^+} \quad . \tag{14.8}$$

Dieses einfache Elektronen-Pol Modell bedeutet eine erhebliche Vereinfachung der korrekten Einteilchen-Green-Funktion. In einer sorgfältigeren Darstellung muß zusätzlich zur Korrektur des Spektrums in (3a) noch die Frequenzabhängigkeit der Quasiteilchen-Energie berücksichtigt werden:

$$E_{\mathbf{nq}} \quad \longrightarrow \quad E_{\mathbf{nq}} + \Sigma_{\mathbf{nq}}(E) \quad , \tag{14.9}$$

wobei für die folgende Diskussion die Selbstenergie an der Stelle $E = E_{\mathbf{nq}}$ der Einfachheit halber auf Null normiert sein soll. Wenn wir von der einfachen linearen Energieabhängigkeit der Selbstenergie in der Nähe der Energie $E_{\mathbf{nq}}$

$$\begin{aligned} \Sigma_{\mathbf{nq}}(E) &= \Sigma_{\mathbf{nq}}(E_{\mathbf{nq}}) + \Sigma'_{\mathbf{nq}} \cdot (E - E_{\mathbf{nq}}) \\ &= \Sigma'_{\mathbf{nq}} \cdot (E - E_{\mathbf{nq}}) \end{aligned} \tag{14.10}$$

ausgehen (vgl. Kap. 2), so ergibt sich die Green-Funktion zu

$$\begin{aligned} G(\mathbf{x},\mathbf{x}',E) &= \sum_{\mathbf{nq}} \psi_{\mathbf{nq}}(\mathbf{x})\psi^*_{\mathbf{nq}}(\mathbf{x}') \cdot \frac{1}{E - E_{\mathbf{nq}} - \Sigma_{\mathbf{nq}}(E) \pm i0^+} \\ &= \sum_{\mathbf{nq}} \psi_{\mathbf{nq}}(\mathbf{x})\psi^*_{\mathbf{nq}}(\mathbf{x}') \cdot \frac{1}{(1 - \Sigma'_{\mathbf{nq}})(E - E_{\mathbf{nq}}) \pm i0^+} \\ &= \sum_{\mathbf{nq}} \psi_{\mathbf{nq}}(\mathbf{x})\psi^*_{\mathbf{nq}}(\mathbf{x}') \cdot Z_{\mathbf{nq}} \frac{1}{E - E_{\mathbf{nq}} \pm i0^+} \quad . \end{aligned} \tag{14.11}$$

Die Energie-Abhängigkeit der Selbstenergie ist generell negativ ($\Sigma' < 0$), und die Renormierungs-Konstante $Z_{\mathbf{nq}} = 1/(1 - \Sigma')$ ist dementsprechend immer kleiner

Selbstenergie und Spektralfunktion des $3d$ Zustands in Ge:

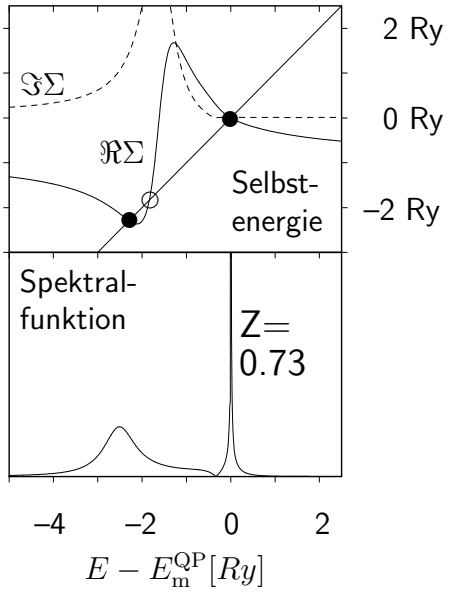


Abbildung 14.1: *Selbstenergie-Funktion und Spektralfunktion des Ge 3d Zustandes im Ge Volumenkristall. Die Selbstenergie-Kurve ist das Ergebnis des Modells (14.14). Die Darstellung erfolgt entlang der reellen Energieachse. In der oberen Figur ist die Energie-Gerade mit eingezeichnet. Ihre Schnittpunkte (•) mit der Selbstenergie definieren die Peaks in der Spektralfunktion. Man beachte, daß der Schnittpunkt bei –2 Ry in einem Energiebereich liegt, in dem der Imaginärteil der Selbstenergie nicht verschwindet. Dadurch wird der zugehörige Satellitenpeak verschoben und verbreitert. Der Schnittpunkt, der mit dem offenen Kreis (∘) markiert ist, trägt zur Spektralfunktion nur sehr wenig bei, weil die Selbstenergie dort sehr steil verläuft.*

als 1. Typische Werte für $Z_{\mathbf{nq}}$ liegen im Bereich zwischen 0.6 und 0.8. Durch die Energie-Abhängigkeit der Selbstenergie bzw. durch die Multiplikation mit der Renormierungskonstante $Z_{\mathbf{nq}}$ weist der Quasiteilchen-Peak in der Spektralfunktion bei der Energie $E_{\mathbf{nq}}$ nicht mehr das volle Gewicht eines freien Teilchens (=1) auf. Statt dessen wird sein Gewicht auf $Z_{\mathbf{nq}} < 1$ reduziert. Hier liegt unmittelbar ein grundsätzliches Problem vor, denn das gesamte Gewicht des Zustands $|\mathbf{nq}\rangle$ in der Spektralfunktion muß Eins betragen:

$$\frac{1}{\pi}\int \left|\Im\left(\frac{1}{E - E_{\mathbf{nq}} - \Sigma_{\mathbf{nq}}(E) \pm i0^{+}}\right)\right| dE = 1 \quad . \tag{14.12}$$

Mit der einfachen, mit $Z_{\mathbf{nq}}$ renormierten Green-Funktion (14.11) ergibt dieses Integral aber nicht 1, sondern nimmt den Wert $Z_{\mathbf{nq}} < 1$ an. Die Green-Funktion hat also nicht das richtige spektrale Gewicht, und die Gesamtzahl der Elektronen wird nicht

erreicht. Das liegt darin begründet, daß die lineare Näherung für die Selbstenergie nur in einem Bereich um die Quasiteilchenenergie $E_{\mathbf{nq}}$ herum gültig ist. Außerhalb dieses Bereiches weicht sie vom linearen Verhalten ab und nimmt eine Struktur an, wie sie aus der oberen Figur in Abb. 14.1 zu entnehmen ist [Hed69, Bec94, Roj95]. Aus der Bestimmungsgleichung

$$E = E_{\mathbf{nq}} + \Sigma_{\mathbf{nq}}(E) \tag{14.13}$$

ergeben sich mit einer so strukturierten Selbstenergie zusätzlich zur Quasiteilchen-Lösung bei $E_{\mathbf{nq}}$ noch weitere Lösungen [Hed69]. Dabei handelt es sich nicht um scharf definierte Peaks, wie sie sich für die Quasiteilchenenergie selbst ergeben, sondern um verbreiterte Strukturen. Der Grund liegt darin, daß der Imaginärteil der Selbstenergie außerhalb des Quasiteilchen-Bereiches nicht verschwindet. Die Energielösungen sind daher komplex, und die zusätzlichen Strukturen in der Spektralfunktion besitzen nur eine endliche Lebensdauer. Sie werden häufig als *Satelliten* bezeichnet.

Der Energieabstand der Satelliten zum Quasiteilchen-Peak liegt meistens in der Größenordnung der Plasmon-Energien. Es stellt sich heraus, daß bei besetzten Quasiteilchen ($E_{\mathbf{nq}} < E_F$) die Satellitenstrukturen vorwiegend bei tieferen Energien auftreten, bei unbesetzten Quasiteilchen hingegen bei höheren Energien [Hed69]. Das läßt sich anschaulich wie folgt verstehen. Wenn ein Elektron aus einem besetzten Zustand ($E_{\mathbf{nq}} < E_F$) aus dem System entfernt werden soll, so muß von außen die Bindungsenergie des Elektrons aufgewendet werden. Wird bei diesem Prozeß ein Plasmon erzeugt, so muß zusätzlich zur Bindungsenergie des Elektrons noch die Plasmon-Energie aufgewendet werden. Bei der Emission des Elektrons können nur Plasmonen erzeugt, nicht aber vernichtet werden, denn bei der betrachteten Temperatur $T = 0$ sind im System keine Plasmon-Zustände besetzt. [Auch bei höheren Temperaturen sind wegen ihrer großen Energie von mehr als 10 eV praktisch keine Plasmonen vorhanden.] Dem Emissionsprozeß ohne Plasmonerzeugung wird man im elektronischen Spektrum die Quasteilchen-Energie $E_{\mathbf{nq}}$ zuordnen, der Emission mit Plasmonerzeugung hingegen eine entsprechend tiefere Bandstruktur-Energie. Wenn hingegen ein Elektron in das System hineingebracht wird ($E_{\mathbf{nq}} > E_F$), so wird seine Bindungsenergie frei. Wenn bei dem Prozeß ein Plasmon erzeugt wird, so erfordert das einen Teil der freiwerdenden Bindungsenergie. Die freiwerdende Bindungsenergie fällt dann entsprechend geringer aus; das entspricht einer *höheren* Bandstrukturenergie als der Quasiteilchen-Peak $E_{\mathbf{nq}}$ ohne Plasmon-Emission.

Bei der hier betrachteten Erzeugung von Plasmonen handelt es sich um intrin-

sische Effekte innerhalb des wechselwirkenden Vielteilchensystems. Darüber hinaus kann ein angeregtes Elektron während des Transports zur Festkörper-Oberfläche, also vor der Messung seiner Energie, in einem extrinsischen Streuprozeß ebenfalls ein Plasmon emittieren und dabei Energie verlieren. Die dadurch erzeugten extrinsischen Satelliten-Strukturen im gemessenen Spektrum lassen sich von den intrinsischen Effekten nur schwer trennen [Woo94].

Wir wollen nun untersuchen, welche Konsequenzen sich aus den Satellitenstrukturen der Green-Funktion für die Selbstenergie und damit letztlich für die Quasiteilchen-Energien ergeben. Hierzu führen wir zwei hintereinander geschaltete GW-Rechnungen durch. Im ersten Schritt (i) bestimmen wir im Rahmen der Standard-GWA die Selbstenergien und deren Energie-Abhängigkeiten für alle Zustände $|n\mathbf{q}\rangle$, die in die Green-Funktion G eingehen. Im zweiten Schritt (ii) berechnen wir mit dieser modifizierten Green-Funktion den Selbstenergie-Operator neu und bestimmen die endgültigen Quasiteilchen-Korrekturen für den Core-Zustand $|m\mathbf{k}\rangle$. Wenden wir uns zunächst der ersten GW-Rechnung (i) zu. Wir stellen für die Selbstenergie ein einfaches Modell auf, das die Behandlung der Energieeffekte erlaubt. Wir gehen von der Tatsache aus, daß die Energie-Abhängigkeit der Selbstenergie – ähnlich wie die Polarisierbarkeit oder die dielektrische Funktion – durch eine Polstruktur gekennzeichnet ist [Roj95]:

$$\Sigma_{\mathrm{n}\mathbf{q}}(E) = a + \sum_{j=1}^{n} \frac{b_j}{E - E_j} \quad . \tag{14.14}$$

Die uns interessierenden Effekte lassen sich schon mit einer zweipoligen Struktur (n=2) beschreiben. Ähnlich wie bei den Plasmon-Pol Modellen für die dielektrische Matrix (siehe Kap. 4) stehen wir vor dem Problem, die Konstante a, die komplexen Pol-Energien E_j und die zugehörigen Gewichte b_j bestimmen zu müssen. Die Bestimmung dieser $2n$+1 Parameter erfordert die Kenntnis der Selbstenergie bei $2n$+1 verschiedenen Testenergien. Diese Selbstenergie-Werte berechnen wir im Rahmen der Standard-GWA. Die Testenergien wählen wir, ausgehend von der reellen Quasiteilchen-Energie $E_{\mathrm{n}\mathbf{q}}$ des Zustands $|\mathrm{n}\mathbf{q}\rangle$, in die komplexe Energie-Ebene hinein. Anschließend passen wir die Parameter des Modells (14.14) an diese berechneten Werte an (siehe Anhang F). Um danach in der zweiten GW-Rechnung (ii) zu einer möglichst einfachen Darstellung zu kommen, berechnen wir aus dem Modell (14.14) gemäß Gl. (14.13) die komplexen Energien $E^{\mathrm{sat}}_{\mathrm{n}\mathbf{q}}$ und Gewichte $Z^{\mathrm{sat}}_{\mathrm{n}\mathbf{q}}$ der Satelliten, die zusätzlich zum Quasiteilchen-Peak (Energie=$E^{\mathrm{QP}}_{\mathrm{n}\mathbf{q}}$, Gewicht=Renormierungskonstante=$Z^{\mathrm{QP}}_{\mathrm{n}\mathbf{q}}$) in der Spektralfunktion auftreten. Die Gewichte ergeben sich dabei

als Renormierungskonstante aus der lokalen Energieabhängigkeit der Selbstenergie an der Satelliten-Energie: $Z = 1/(1 - \Sigma')$. Wir nehmen der Einfachheit halber an, daß die räumlichen Eigenschaften der Satelliten durch die Quasiteilchen-Amplituden $\psi_{\mathbf{nq}}(\mathbf{r})$ gegeben sind. In Abb. 14.1 stellen wir am Beispiel des Ge $3d$ Zustands in Ge einen typischen Energieverlauf der Selbstenergie und die zugehörige spektrale Dichte entlang der reellen Energieachse dar, wie sie sich aus unserem Zugang ergibt.

Wir haben diese Erweiterung der Green-Funktion in der zweiten GW-Rechnung (ii) berücksichtigt. Dadurch ändert sich die Selbstenergie des Zustands $|m\mathbf{k}\rangle$. Der Austausch-Anteil zur Selbstenergie bleibt unverändert, denn er hängt von der Energiestruktur der Green-Funktion gar nicht ab. Der Korrelationsanteil hingegen hängt von der energetischen Struktur der Spektralfunktion ab. Wir können zunächst in einer vereinfachten Überlegung den zu erwartenden Effekt für die Korrelationsenergie und die Bandstruktur-Energie des Core-Zustands grob abschätzen. Wenn wir annehmen, daß in der Standard-GWA (2) die Korrelationsenergie des betrachteten Zustands (Quasiteilchen-Energie $E_{\rm m}^{\rm QP}$) als Kopplung mit sich selbst und mit einem Plasmon der Energie ω_P beschrieben werden kann, so ergibt das in (14.6) angegebene Energie-Integral für die herkömmliche Rechnung ohne Renormierung des Spektrums einen Wert

$$A \cdot \frac{1}{E_{\rm m}^{\rm QP} - E_{\rm m}^{\rm QP} + \omega_P} = A \cdot \frac{1}{\omega_P} \quad , \tag{14.15}$$

der als Faktor in die Korrelationsenergie eingeht. Durch die Berücksichtigung der Renormierung und der Satelliten-Struktur verändert sich dieser Faktor. Wenn man annimmt, daß der Quasiteilchen-Peak nur noch das Gewicht $Z < 1$ hat und das restliche Gewicht (1–Z) einem Satelliten bei der tieferen Energie $E_{\rm m}^{\rm QP} - \Delta^{\rm sat}$ zugeordnet werden kann, so ändert sich die Green-Funktion des Zustands wie folgt:

$$\frac{1}{E - E_{\rm m}^{\rm QP}} \quad \longrightarrow \quad Z\frac{1}{E - E_{\rm m}^{\rm QP}} + (1 - Z)\frac{1}{E - E_{\rm m}^{\rm QP} + \Delta^{\rm sat}} \quad . \tag{14.16}$$

Daraus resultiert eine Änderung des obigen Integrals (14.15) zu

$$\begin{aligned} & A \cdot \left(\frac{Z}{E_{\rm m}^{\rm QP} - E_{\rm m}^{\rm QP} + \omega_P} + \frac{1 - Z}{E_{\rm m}^{\rm QP} - E_{\rm m}^{\rm QP} + \Delta^{\rm sat} + \omega_P} \right) \\ = \quad & A \cdot \frac{1}{\omega_P} \cdot \left(Z + (1 - Z)\frac{\omega_P}{\omega_P + \Delta^{\rm sat}} \right) \quad . \end{aligned} \tag{14.17}$$

Wenn wir annehmen, daß die Renormierungskonstante Z=0.6 beträgt und die Energieabsenkung $\Delta^{\rm sat}$ so groß ist wie die Plasmon-Energie ω_P, dann wird durch den

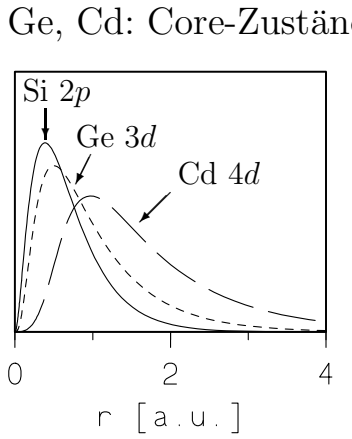


Abbildung 14.2: *Atomare Pseudo-Wellenfunktionen* $r\psi(\mathbf{r})$ *der Core-Zustände Si 2p, Ge 3d und Cd 4d.*

Faktor $(Z + (1 - Z)\omega_P/(\omega_P + \Delta^{\text{sat}})) = 0.8$ die Korrelationsenergie um 20 % reduziert. Die Korrelationsenergie eines Core-Elektrons liegt in der Größenordnung +10 eV; sie ist positiv. Durch die Reduktion dieses Energiebeitrags zur Selbstenergie erwarten wir folglich gegenüber (3a) eine weitere Absenkung der Bandstrukturenergie in der Größenordnung einiger eV.

In unseren konkreten, vollständigen Rechnungen im Rahmen des Modells (14.14) stellen wir tatsächlich derartige Energieabsenkungen fest (siehe Spalte 4 in Tabelle 14.1). Sie betragen –2.5 eV für den Zustand Si 2*p*, –1.8 eV für Ge 3*d* sowie –0.9 eV für Cd 4*d*. Auch diese Absenkungen sind für den tieferliegenden, stark lokalisierten Core-Zustand Si 2*p* größer als für den weniger stark lokalisierten Semicore-Zustand Cd 4*d*, der höhere Valenzbandenergien aufweist.

Bei den ersten drei Näherungen (1, 2, 3a) sind die Abweichungen der jeweiligen Core-Energien vom experimentellen Wert für den Si 2*p* Zustand am größten und für Cd 4*d* am kleinsten. Für die *Korrekturen* durch die einzelnen Näherungen (1→2, 2→3a, 3a→3a+b) gilt der gleiche Trend. Diese Trends müssen im Zusammenhang mit der Lokalisierung und der Bindungsenergie der Zustände betrachtet werden. Wir vergleichen in Abb. 14.2 die radialen Wellenfunktionen der drei Zustände miteinander. Der Si 2*p* Zustand ist am stärksten lokalisiert und weist die steilste Wellenfunktion auf, während Cd 4*d* eher moderat lokalisiert ist. Gleichzeitig hat Si 2*p* eine wesentlich stärkere Bindungsenergie als Cd 4*d*. Wir stellen daher fest, daß die systematische Unterschätzung der Bindungsenergie eines Core-Zustands durch die

Zugänge 1, 2 und 3a desto größer ausfällt, je stärker der Zustand lokalisiert und je größer seine Bindungsenergie ist.

Die Ergebnisse, die wir mit dem endgültigen Zugang 3a+b erzielt haben, stimmen bis auf 0.4 eV mit dem Experiment überein. Der relative Fehler der Bindungsenergie beträgt maximal 1 %, während die LDA Fehler von bis zu 20 % aufweist. Das suggeriert eine deutlich gesteigerte quantitative Aussagefähigkeit der Methoden durch die Aktualisierung des Spektrums und die Berücksichtigung der Satelliten. Dieser Eindruck sollte etwas relativiert werden. Zum einen wurde die Satellitenstruktur nicht exakt, sondern lediglich im Rahmen des einfachen Modells (14.14) berücksichtigt. Wir verfeinern das Modell nicht weiter, da es uns primär darauf ankommt, generelle Trends und Mechanismen aufzuzeigen. Zum anderen beeinflussen die Satellitenstrukturen nicht nur die Core-Energien, sondern auch die Valenz- und Leitungsbänder. Plasmon-induzierte Satelliten finden sich unterhalb der Valenzbänder und oberhalb der Leitungsbänder [Hed69]. Analog zu den Core-Zuständen bewirken sie auch hier eine Abschwächung der Korrelationsenergien der Valenz- und Leitungszustände. Diese sind für Valenzzustände positiv und für Leitungsbandzustände negativ (vgl. Abb.5.1). Ihre Reduktion bewirkt somit eine energetische Absenkung der Valenzbänder und Anhebung der Leitungsbänder. Dadurch wird das Gap zwischen ihnen erheblich vergrößert. Wir erhalten im Rahmen der GWA (3a+b) Gapenergien von 1.49 eV für Si, 0.93 eV für Ge und 3.21 eV für CdS, die deutlich über den experimentellen Werten (1.17 eV, 0.74 eV bzw. 2.48 eV) liegen. Ähnliche Effekte wurden von de Groot et al. [Gro95] bei Untersuchungen an einem quasieindimensionalen Modell-Halbleiter festgestellt. Die Autoren führen eine selbstkonsistente GW-Rechnung ($\Sigma = GW$) durch, die die komplette Energieabhängigkeit von G mit allen Satellitenstrukturen berücksichtigt. Die damit erzielte selbstkonsistente Gapenergie ist weit größer als die Gapenergie des ersten Iterationsschritts ($\Sigma^0 = G^0W^0$), der unserer Standard-GWA entspricht.

Unsere Ergebnisse scheinen anzudeuten, daß zur quantitativ richtigen Beschreibung der Core-Elektronen im Rahmen der GWA die Satellitenstruktur berücksichtigt werden sollte. Die Bestimmung der Valenzband- und Leitungsbandstruktur liefert hingegen im Rahmen der Standard-GWA, also ohne Satelliten, eine bessere Übereinstimmung mit experimentellen Daten. Darin liegt eine gewisse Willkür verborgen, die im Rahmen eines ab-initio Zugangs nicht befriedigen kann. Die Ursache liegt letztlich darin begründet, daß die GWA doch eine relativ einfache Näherung für die Selbstenergie darstellt. Es ist generell umstritten, ob die Satelliten-

Strukturen überhaupt in der Green-Funktion implementiert werden sollten. Möglicherweise stellt der GW-Selbstenergie-Operator mit der nichtrenormierten Green-Funktion eine prinzipiell bessere Näherung für den exakten Selbstenergie-Operator (inklusive aller Vertex-Beiträge) dar als ein Selbstenergie-Operator, der die Satellitenstrukturen enthält [Blo72, Bec94].

Kapitel 15

Core-Niveau Verschiebungen an Oberflächen

Die Untersuchung von Core-Niveaus spielt bei der experimentellen Analyse der Zusammensetzung und Struktur eines Festkörpers eine wichtige Rolle. Im Gegensatz zur delokalisierten Valenzelektronenstruktur können die Core-Niveaus auch im Festkörper einzelnen Atomen zugeordnet werden. Je nach Element weisen die Core-Niveaus charakteristische Bindungsenergien auf. Wenn im Photoemissionsspektrum die entsprechende Linie vorliegt, so kann daraus eindeutig auf das Vorhandensein des zugehörigen Elements in der Probe geschlossen werden. Die quantitative Untersuchung der Emissionsintensitäten erlaubt Aussagen über die Konzentration der Elemente in der Probe und über ihre stöchiometrische Zusammensetzung.

Die genaue Bindungsenergie eines Core-Elektrons hängt davon ab, in welcher chemischen Umgebung sich das Atom befindet. Durch Ladungsverschiebungen und Änderungen der dielektrischen Abschirmung kommt es zu Verschiebungen des Core-Niveaus gegenüber der Bindungsenergie im freien Atom oder im Volumenkristall. Die Untersuchung dieser Verschiebungen gestattet detaillierte Aussagen über die mikroskopische Struktur der Probe. Hierzu zählen neben der reinen Verschiebung des Core-Niveaus auch seine Aufspaltung durch eine reduzierte Symmetrie der Umgebung des Atoms. Die Core-Niveau Spektroskopie läßt sich auf mehrere Arten zur Analyse der mikroskopischen Struktur eines Festkörpers einsetzen:
(1) Die Untersuchung der Core-Elektronen des Substrat-Materials erlaubt Aussagen über Störungen des Kristalls durch Gitterfehler, Defekte, Grenz- und Oberflächen. Neben der immer vorhandenen Volumen-Linie treten weitere Emissionslinien auf, die den Substrat-Atomen nahe der Störung zuzuordnen sind.

(2) Eine weitere Bedeutung der Core-Niveaus des Substrats liegt in der Verwendung als Referenzenergie, etwa für Untersuchungen an Grenz- und Oberflächen. Der energetische Abstand des Volumen-Core-Niveaus zum VBM ändert sich durch die Modifikation der Grenz- und Oberflächen nicht. Die Messung des Core-Niveaus relativ zum Ferminiveau erlaubt die genaue Bestimmung des Abstandes zwischen VBM und Ferminiveau. Daraus lassen sich Aussagen über Barrierenhöhen und Valenzband-Offsets gewinnen.
(3) Wenn im Festkörper Fremdatome vorhanden sind, so erlaubt die Analyse ihrer Core-Niveaus Aussagen über ihre chemische Umgebung und ihre genaue Position im Substrat-Material.

Wir behandeln in diesem Kapitel den ersten dieser Punkte und untersuchen die Verschiebungen des Si $2p$ Niveaus an freien und adsorbat-bedeckten Si Oberflächen gegenüber dem Volumen-Niveau. Dieses Problem wurde mehrfach im Rahmen der LDA behandelt [Peh93, Bla94b, Pas96]. Dabei kamen zwei unterschiedliche Ansätze zur Anwendung:
(i) Das effektive LDA-Potential in der Umgebung der Atomkerne unterscheidet sich an der Oberfläche vom Volumen. Die Differenz bewirkt eine Core-Niveau Verschiebung.
(ii) Der Ansatz (i) betrachtet nur den Anfangszustand des Photoemissionsprozesses. Dabei bleibt unberücksichtigt, daß auch der Endzustand (=das Core-Loch) sich für Emission aus Oberflächenatomen und aus dem Volumen unterscheidet. Einen Zugang hierzu kann man durch Berechnungen der totalen Energien des Systems innerhalb der LDA erhalten [Peh93]. Die Erzeugung eines Core-Loches kann durch die Verwendung eines Si Pseudopotentials (PP) simuliert werden, bei dessen Konstruktion der atomare Si $2p$ Zustand nur mit fünf statt sechs Elektronen besetzt wird. In zwei separaten LDA Rechnungen für das Oberflächensystem wird dieses PP anstelle des herkömmlichen Si PP einmal im Volumen, in der anderen Rechnung am Oberflächenatom plaziert. Die Differenz der totalen Energien der beiden Rechnungen liefert dann die Core-Niveau Verschiebung am Oberflächenatom gegenüber dem Volumenwert.

Die beiden Vorgehensweisen (i) und (ii) innerhalb der LDA weisen zwei systematische Schwächen auf. Die eine besteht in der Verwendung der LDA zur Beschreibung der Austausch-Korrelations-Effekte. Es ist a priori nicht zu erkennen, ob die LDA in der Lage ist, die Core-Niveaus in der Nähe von Oberflächen genau genug zu beschreiben. Der andere Schwachpunkt besteht darin, daß der zu untersuchende

Core-Zustand in den Rechnungen gar nicht explizit behandelt wird. Er wird statt dessen in Form des speziellen Pseudopotentials indirekt berücksichtigt. Das erleichtert zwar die numerische Durchführbarkeit der Rechnungen erheblich, da nur glatte Pseudopotentiale und Wellenfunktionen auftreten. Dennoch wäre es wünschenswert, den fraglichen Core-Zustand explizit in die Rechnung zu implementieren, um seine Wechselwirkung mit der jeweiligen Umgebung direkt bestimmen zu können.

Unsere Rechenmethoden erlauben es uns, beide Problempunkte zugleich anzusprechen. Im vorangegangen Kapitel hatten wir explizit die Quasiteilchen-Energien von Core-Zuständen in Volumenkristallen bestimmt. Dabei wurden die Zustände direkt in Form von Valenzzuständen behandelt. Nun wenden wir denselben Formalismus auf Slabsysteme an und bestimmen die Quasiteilchen-Energien der Core-Zustände an der Oberfläche und in der volumenartigen Slabmitte. Die Energien und ihre Differenz lassen sich direkt aus der Quasiteilchen-Bandstruktur ablesen. Wir beschränken uns dabei auf den GW-Formalismus 3a, in dem die Satelliten-Struktur der Green-Funktion noch keine Beachtung findet. In einer Testrechnung haben wir festgestellt, daß sich durch Berücksichtigung der Satelliten zwar die absolute Energie der Core-Zustände ändert, nicht aber die Core-Niveau Verschiebungen zwischen Volumen und Oberfläche.

Bevor wir die einzelnen Si Oberflächen diskutieren, merken wir noch an, daß wir in unseren Rechnungen die Spin-Bahn Aufspaltung der Si $2p$ Zustände von 0.61 eV zunächst nicht berücksichtigen. Wenn nicht anders erwähnt, gelten die Core-Niveau Verschiebungen für die beiden Zustände $2p_{1/2}$ und $2p_{3/2}$ gleichermaßen.

15.1 Die As:Si(111)-(1×1) Oberfläche

Wir untersuchen zunächst die As:Si(111)-(1×1) Oberfläche. Anhand dieses Beispiels werden die wesentlichen Elemente unserer Vorgehensweise erläutert. Unser Zugang beruht auf zwei grundlegenden Annahmen:

- Wenn ein einzelnes Atom eine besondere chemische Umgebung besitzt, die es von den anderen Atomen im Substrat unterscheidet, so bildet sich an ihm ein Core-Zustand aus, der mangels energetischer Resonanz nicht mit Core-Zuständen an anderen Atomen hybridisiert. Diese Eigenschaft ist grundsätzlich immer erfüllt.

- Die Berechnung der Struktur der Valenz-Elektronen, die für die chemische

Umgebung des fraglichen Atoms verantwortlich ist, hängt nicht davon ab, ob die Core-Elektronen in technischer Hinsicht als Core- oder als Valenzelektronen behandelt werden.

Im vorliegenden Falle der Si Oberflächen ist die zweite Annahme sehr gut erfüllt (siehe Abschnitt 13.4). Rechnungen mit dem Si^{4+} PP und mit dem Si^{12+} PP ergeben fast exakt gleiche Bandstrukturen. Konkret bedeutet das, daß es für die Berechnung der Valenz-Eigenschaften an einem Atom irrelevant ist, ob an benachbarten Atomen Si^{4+} PP'e oder Si^{12+} PP'e verwendet werden. Im folgenden sollen gezielt die Bindungsenergien des Si $2p$ Zustands an einzelnen Si Atomen an Oberflächen untersucht werden. Das geschieht dadurch, daß zur Beschreibung des fraglichen Atoms das Si^{12+} PP eingesetzt wird, das den Si $2p$ Zustand als Valenzzustand beinhaltet. In einer Berechnung der elektronischen Bandstruktur erhalten wir dann automatisch die Energie dieses Zustands am fraglichen Atom. Die obigen Annahmen erlauben es, ohne Verfälschung der Ergebnisse das Si^{12+} PP nur an denjenigen Si Atomen einzusetzen, an denen wir die Core-Niveaus bestimmen wollen. Das sind einerseits diejenigen Si Atome an der Oberfläche, an denen das $2p$ Niveau eine Verschiebung erfährt, und ein Referenzatom in der volumenartigen Slabmitte, dessen $2p$ Niveau wir zur Festlegung der Referenzenergie und zur Ermittlung der Verschiebung benötigen. In unseren Rechnungen beschreiben wir alle anderen Si Atome mit dem Si^{4+} PP, um den ohnehin enormen numerischen Aufwand zu begrenzen. Innerhalb der LDA haben wir vergleichende Rechnungen mit mehr Si^{12+} PP'en durchgeführt und dabei keine Veränderung der fraglichen Core-Niveaus festgestellt. Die As Atome beschreiben wir in allen Rechnungen mit einem konventionellen As^{5+} PP [Bac82]. Die Core-Niveaus der As Atome werden von uns nicht betrachtet.

Wir stellen die Geometrie der As:Si(111)-(1×1) Oberfläche in Abb. 15.1 nochmals vor; sie ist mit der in Abschnitt 9.4 untersuchten identisch. Im experimentellen Si $2p$ Core-Niveau Spektrum dieser Oberfläche wird neben der Volumen-Linie nur ein weiteres Niveau beobachtet, das gegenüber der Volumen-Linie um 0.75 eV zu niedrigeren Bandstruktur-Energien abgesenkt ist [Uhr87, Bul96]. Es liegt nahe, dieses Niveau den Si $2p$ Zuständen an den Si Atomen in der zweiten Schicht, die chemische Bindungen mit den As Atomen ausbilden, zuzuordnen. An diesen Atomen ist wegen der Nähe zur Oberfläche und wegen der Bindungen zu den elektronegativeren As Atomen die stärkste Core-Niveau Verschiebung zu erwarten. Die chemische Umgebung der tiefer im Substrat liegenden Si Atome ist der chemischen Umgebung im

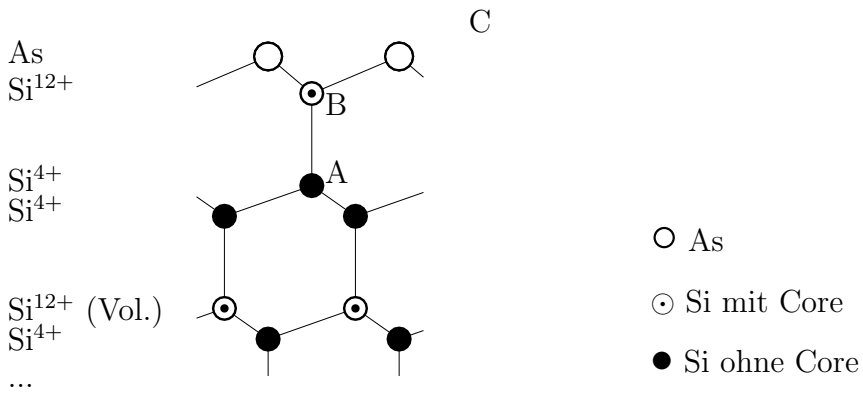


Abbildung 15.1: *Struktur des Slabs zur Untersuchung der Core-Niveau Verschiebung an der As:Si(111)-(1×1) Oberfläche (vgl. Kap. 9). Die offenen Kreise (○) markieren As-Adatome, die fetten Kreise (●) Si-Atome. Die Kreise mit dem Punkt in der Mitte (⊙) markieren Si Atome mit Si^{12+} PP, an denen die Core-Niveaus ermittelt werden. Die Raumpunkte A, B und C definieren die Richtungen, entlang derer im folgenden die Ladungsdichte und das Potential dargestellt werden.*

Volumen schon so ähnlich, daß ihre Core-Niveau Verschiebung zum Volumen sehr gering ausfällt und experimentell nicht mehr auflösbar ist.

Wir berechnen die Core-Niveau Verschiebung an den Si Atomen der zweiten Schicht gegenüber dem Volumen mit unserem Zugang. Dazu plazieren wir am fraglichen Oberflächenatom sowie an einem Atom im Bulk das Si^{12+} PP (in Abb. 15.1 als ⊙ gekennzeichnet). Ein großer Beitrag zur Core-Niveau Verschiebung kommt durch die Modifikation der Valenz-Ladungsdichte und die Veränderung des klassischen elektrostatischen Potentials am Ort des fraglichen Si Atoms zustande. Anders als im Volumen ist es nicht von vier Si Atomen, sondern von einem Si Atom und drei As Atomen umgeben. Jenseits der As Atome beginnt zudem das Vakuum. Wir stellen in Abb. 15.2 die Valenzladungsdichte am fraglichen Si Atom entlang der rückseitigen Si-Si Bindung und entlang der Si-As Bindung mit den As Oberflächenatomen dar (durchgezogene Linien). Die gestrichelten Linien geben die entsprechende Ladungsdichte an einem Si Atom im Volumen entlang der Bindungsrichtungen an. Aufgrund der stärkeren Elektronegativität von As im Vergleich zu Si wird entlang der Si-As Bindung elektronische Ladung zum As Atom verschoben. Daraus ergibt sich am Si Atom eine leichte positive Partialladung. Das elektrostatische Potential steigt dadurch an dem Atom leicht an, und die Elektronen an dem Atom erfahren wegen

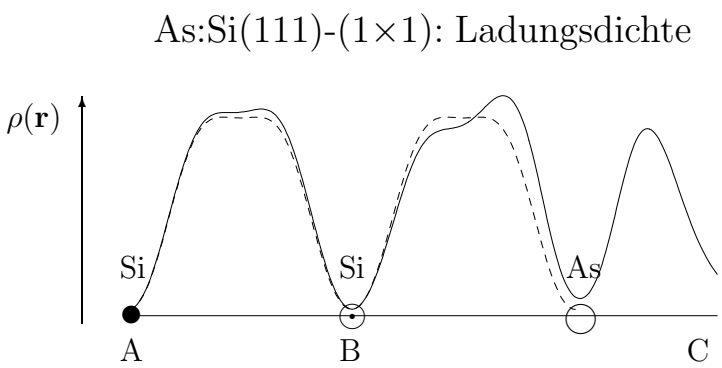


Abbildung 15.2: *Ladungsdichte am Sublayer Si Atom ⊙ an der As:Si(111)-(1×1) Oberfläche. Die gestrichelte Kurve markiert die Ladungsdichte entlang der entsprechenden Bindungen im Volumen.*

ihrer negativen Ladung eine Potential*absenkung*. Sie beträgt in der Umgebung des Atomkerns –0.9 eV. Wir stellen das Potential der Si $2p$ Elektronen in Abb. 15.3 entlang der Bindungsrichtungen dar. In der unteren Figur ist die Differenz zwischen den Potentialen am Oberflächenatom und an einem Volumen-Referenzatom vergrößert dargestellt.

Das unterschiedliche Potential am Oberflächenatom führt – zusammen mit einer minimalen Differenz der kinetischen Energien – zu einer Absenkung des dortigen Si $2p$ Niveaus um zunächst –0.94 eV gegenüber dem Niveau am Referenzatom in der Slabmitte. Dies ist aber nicht der einzige Beitrag zur Core-Niveau Verschiebung. Die Selbstenergie der Core-Zustände unterscheidet sich ebenfalls am Oberflächenatom und in der Slabmitte. Unsere Methoden bieten zwei Möglichkeiten, die Selbstenergie direkt zu bestimmen. Die Bestimmung der XC-Energie im Rahmen der LDA weist systematische Schwächen auf, die wir im letzten Kapitel dargelegt haben. Man könnte dennoch vermuten, daß die LDA die Core-Niveau *Verschiebungen* zwischen Volumen- und Oberflächenatomen richtig beschreibt. Prinzipiell ist das jedoch nicht der Fall [Peh93]. In unserer Rechnung beträgt der Unterschied der LDA XC-Energien des Si $2p$ Zustands am Si Sublayeratom und im Volumen allerdings lediglich 0.01 eV. Die Core-Niveau Verschiebung zum Volumen ändert sich dadurch nur unwesentlich auf –0.93 eV. Die Ursache ist darin zu sehen, daß das lokale LDA XC-Potential die Information über die Oberfläche nicht enthält. Obwohl der zu untersuchende Si $2p$ Zustand nur 1 Å von der Oberflächenebene entfernt ist, weicht die Ladungsdichte-Verteilung an dem Si Atom nur minimal von der im Volumen ab (vgl. Abb. 15.2). Dementsprechend ist auch das LDA XC Potential fast mit dem im

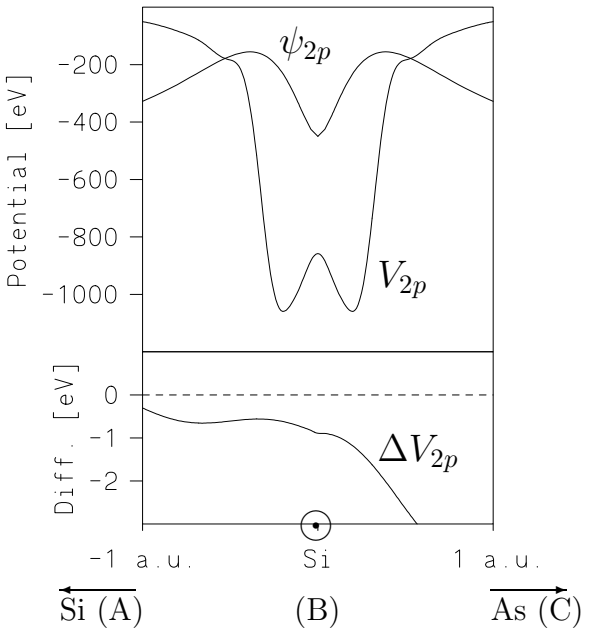


Abbildung 15.3: *Oben: Potential der Si 2p Elektronen am Si Sublayeratom (⊙) an der As:Si(111)-(1×1) Oberfläche. Die Darstellung erfolgt entlang der Bindungsrichtungen AB zum nächsten Si Atom sowie BC zum As Adatom (vgl. Abb. 15.1). Der Vollständigkeit halber ist die Wellenfunktion von Si 2p mit aufgeführt. Unten: Differenz des Potentials zum entsprechenden Potential im Si Volumenkristall. Man beachte die vergrößerte Skala.*

Volumen identisch. In der Realität aber weichen die Abschirm-Eigenschaften des Systems an der Oberfläche deutlich vom Volumen ab. Sie führen zu einer Veränderung der Selbstenergie des Si $2p$ Zustands gegenüber dem Volumen [Peh93].

Unser GW-Zugang erlaubt die direkte Bestimmung der Selbstenergie und der Quasiteilchen-Energie des Core-Zustands unter Einschluß aller Abschirm- und Oberflächeneffekte. Die von uns berechnete Selbstenergie des $2p$ Zustands liegt an der Oberfläche um 0.06 eV über der im Volumen. Dieser Beitrag zur Core-Niveau Verschiebung ist größer als der entsprechende Beitrag der LDA. Er fällt dennoch erstaunlich gering aus. Das liegt daran, daß das Si Sublayeratom nahezu tetraedrisch von As und Si Atomen umgeben ist und somit eine ähnliche Konfiguration wie im Volumen besitzt. Der relativ kurzreichweitige Selbstenergie-Operator ergibt dementsprechend auch eine ähnliche Selbstenergie wie im Volumen.

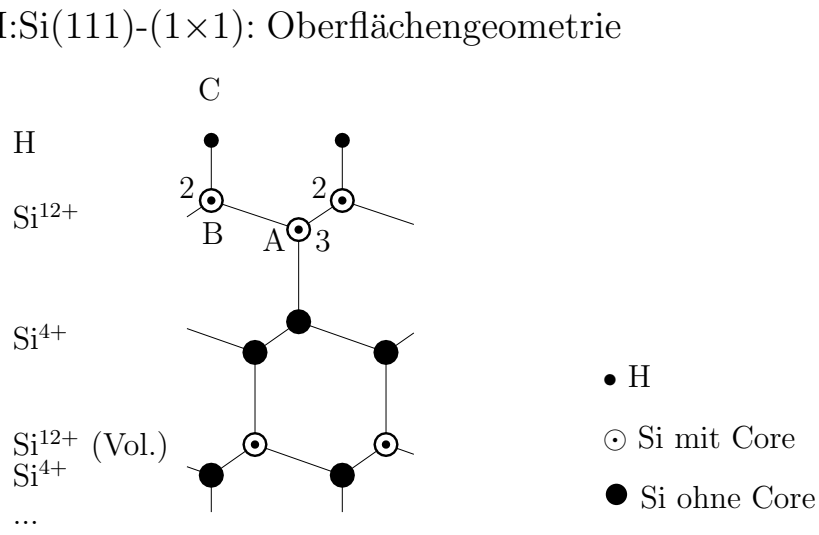


Abbildung 15.4: *Struktur des Slabs zur Untersuchung der Core-Niveau Verschiebungen an der H:Si(111)-(1×1) Oberfläche (vgl. Kap. 9). Die kleinen Punkte (•) markieren H-Adatome, die großen Punkte (●) Si-Atome. Die Kreise mit dem Punkt in der Mitte (⊙) markieren Si Atome mit Si^{12+} PP, an denen die Core-Niveaus ermittelt werden. Die Raumpunkte A, B und C definieren die Richtungen, entlang derer im folgenden die Ladungsdichte und das Potential dargestellt werden.*

Insgesamt finden wir für das Si Sublayeratom eine Si $2p$ Core-Niveauverschiebung von –0.88 eV. Sie stimmt gut mit dem experimentellen Befund von –0.75 eV überein [Uhr87, Bul96]. Wir merken an, daß wir in unseren Rechnungen, in denen wir die Spin-Bahn Wechselwirkung vernachlässigt haben, sowohl in der LDA als auch in der GWA eine energetische Entartung der drei Si $2p$ Zustände an den Si Sublayeratomen beobachten. Eine Kristallfeld-Aufspaltung durch die geänderte Symmetrie an der Oberfläche tritt nicht auf. Das bedeutet, daß im endgültigen Spektrum sowohl der Si $2p_{1/2}$ Zustand als auch der zweifach entartete Si $2p_{3/2}$ Zustand die gleiche Verschiebung von –0.88 eV erfahren.

15.2 Die H:Si(111)-(1×1) Oberfläche

Wir hatten in Abschnitt 9.2 die Valenzband-Struktur der H-terminierten H:Si(111)-(1×1) Oberfläche untersucht. Hier greifen wir das System wieder auf und bestimmen die Si $2p$ Core-Niveau Verschiebungen an den oberflächennahen Si Atomen der zweiten und dritten Schicht. Dabei gehen wir genau wie bei der As:Si(111)-(1×1)

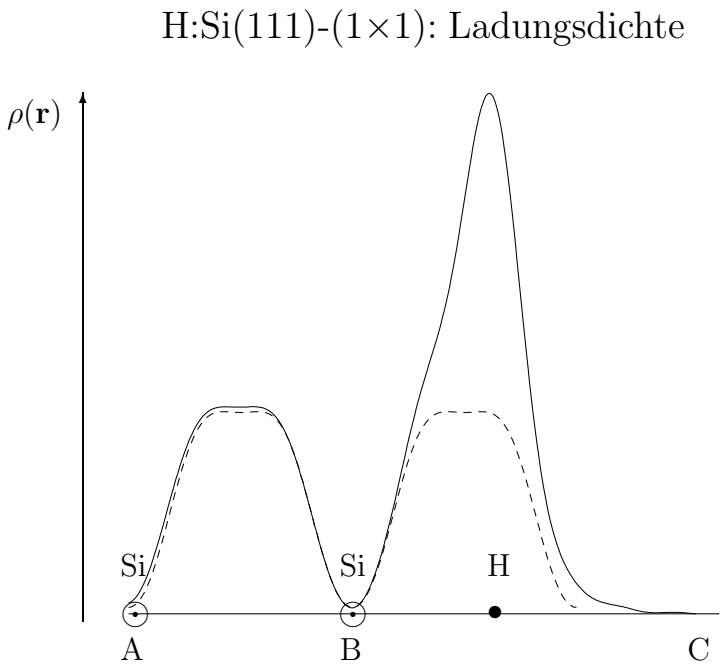


Abbildung 15.5: *Ladungsdichte am Sublayer Si Atom an der H:Si(111)-(1×1) Oberfläche. Die gestrichelte Kurve markiert die entsprechende Ladungsdichte entlang der Bindungen im Volumen.*

Oberfläche vor. [In unseren Rechnungen haben wir beide Oberflächen am gleichen Slabsystem untersucht. Die eine Slab-Oberfläche war As-terminiert, die andere H-terminiert.] In Abb. 15.4 ist die Oberfläche nochmals dargestellt.

In experimentellen Untersuchungen des Si $2p$ Core-Niveau Spektrums der Oberfläche werden mehrere oberflächeninduzierte Emissionslinien beobachtet [Bla94b, Dum96, Hor96], die den Si Atomen der zweiten und der dritten Schicht zuzuordnen sind. Wir berechnen im Folgenden an diesen Si Atomen die Verschiebungen zu einem Volumen-Referenzatom. Diese Atome sind in Abb. 15.4 als $\odot$ gekennzeichnet. Wir untersuchen zunächst die Ladungsdichte und das Potential an den Si Atomen der zweiten Schicht. In Abb. 15.5 stellen wir die Ladungsdichte entlang der Bindungen AB zu den Si Atomen der dritten Schicht sowie entlang der Bindung BC zu den H Adatomen dar. Es werden zwei Unterschiede zur Ladungsdichte an der As:Si(111)-(1×1) Oberfläche deutlich. Zum einen weicht an den Si Atomen der dritten Schicht (A; auf der linken Seite der Abbildung) die Ladungsdichte ein wenig von der Volumen-Ladungsdichte ab. Anders als bei As:Si(111)-(1×1) wird nun auch dieses Atom eine Core-Niveau Verschiebung aufweisen. Zum anderen unterscheidet sich die Ladungsdichte entlang der Si-H Bindung deutlich von der entlang der Si-

As Bindung. Das H-Adatom befindet sich in geringem Abstand vom Si Atom (1.52 Å). Obwohl es nur ein einzelnes Elektron im stark lokalisierten $1s$ Zustand besitzt, erhöht sich die Ladungsdichte am Si Atom im Vergleich zum Volumenwert. Das scheint im Widerspruch dazu zu stehen, daß H eine größere Elektronegativität besitzt als Si. Diese größere Elektronegativität bedeutet aber nur, daß die Ladung auf der Bindung im Vergleich zur Summe der *atomaren* Ladungsdichten verschoben wird. Über die Ladungsdichten in chemisch unterschiedlichen Bindungen sagt die Elektronegativität nur wenig aus.

Die elektronische Ladungsverteilung in der Umgebung des Si Sublayer-Atoms weicht von der idealen Tetraedersymmetrie ab. Das führt dazu, daß (im Gegensatz zur oben untersuchten As:Si(111)-(1×1) Oberfläche) die Entartung der Si $2p$ Zustände an diesem Si Atom aufgehoben wird und die Niveaus aufspalten. Die folgende Diskussion gestaltet sich dadurch schwieriger als bei der As:Si(111)-(1×1) Oberfläche. Erschwerend kommt hinzu, daß die Kristallfeld-Aufspaltung im realen Core-Niveau Spektrum mit der Spin-Bahn Aufspaltung zusammen betrachtet werden muß, die aber in unseren Rechnungen zunächst nicht enthalten ist.

Wir führen die folgende Analyse in mehreren Schritten durch. Wir vernachlässigen zunächst die Spin-Bahn Wechselwirkung. Die Oberfläche weist im Vergleich zum Volumen nur noch C_{3v} Symmetrie auf. Die drei im Atom und im Volumenkristall entarteten Si $2p$ Zustände spalten daher an jedem oberflächennahen Atom i in Zustände auf, deren Winkelverteilung durch die Winkelfunktionen $Y_{l,m}$ mit l=1, $m = -1$, 0, und 1 gegeben sind [Bla94b]. Die z-Achse ist dabei durch die Oberflächen-Normale festgelegt. Die zugehörigen Energieniveaus $E^{(i)}_{1,m}$ treten direkt in den von uns berechneten Bandstrukturen auf. Als Folge der C_{3v} Symmetrie gilt $E^{(i)}_{1,-1}$=$E^{(i)}_{1,1}$. Wir beobachten also an jedem Oberflächenatom zwei Energiewerte, von denen einer doppelt auftritt. Aus diesen Energien erhalten wir zunächst die Core-Niveau Verschiebungen $\Delta E^{(i)}_{1,m} = E^{(i)}_{1,m} - E^{\mathrm{Vol}}$. Diese vorläufigen Core-Niveau Verschiebungen lassen sich formal als Erwartungswerte eines Differenz-Hamiltonians $\Delta H^{(i)}$ am Atom i zum Volumenkristall ausdrücken:

$$\Delta E^{(i)}_{l,m} = \langle l,m|\Delta H^{(i)}|l,m\rangle \quad . \tag{15.1}$$

Aufgrund der Spin-Bahn Wechselwirkung liegen die Zustände $\{|l,m\rangle\}$ im realen System aber gar nicht als Eigenzustände vor, sondern werden mit den Spin-Funktionen $\{|s,m_s\rangle\}$ zu den folgenden Zuständen $\{|j,m_j\rangle\}$ gekoppelt:

$$|3/2,\pm 3/2\rangle \;=\; |1,\pm 1\rangle\,|1/2,\pm 1/2\rangle \tag{15.2}$$

$$|3/2,\pm 1/2\rangle = \sqrt{2/3}\,|1,0\rangle\,|1/2,\pm 1/2\rangle + \sqrt{1/3}\,|1,\pm 1\rangle\,|1/2,\mp 1/2\rangle \quad (15.3)$$

$$|1/2,\pm 1/2\rangle = \pm\sqrt{2/3}\,|1,\pm 1\rangle\,|1/2,\mp 1/2\rangle \mp \sqrt{1/3}\,|1,0\rangle\,|1/2,\pm 1/2\rangle \quad . (15.4)$$

Die Core-Niveau Veschiebungen $\Delta E^{(i)}_{j,m_j}$ dieser neuen, realen Eigenzustände berechnen sich ebenfalls als Erwartungswerte des Differenz-Hamiltonians $\Delta H^{(i)}$:

$$\Delta E^{(i)}_{j,m_j} = \langle j, m_j|\Delta H^{(i)}|j, m_j\rangle \quad . \quad (15.5)$$

Unter Ausnutzung der bereits vorliegenden Verschiebungen $\Delta E^{(i)}_{1,m}$ ergibt sich:

$$\Delta E^{(i)}_{3/2,\pm 3/2} = \Delta E^{(i)}_{1,\pm 1} \quad (15.6)$$

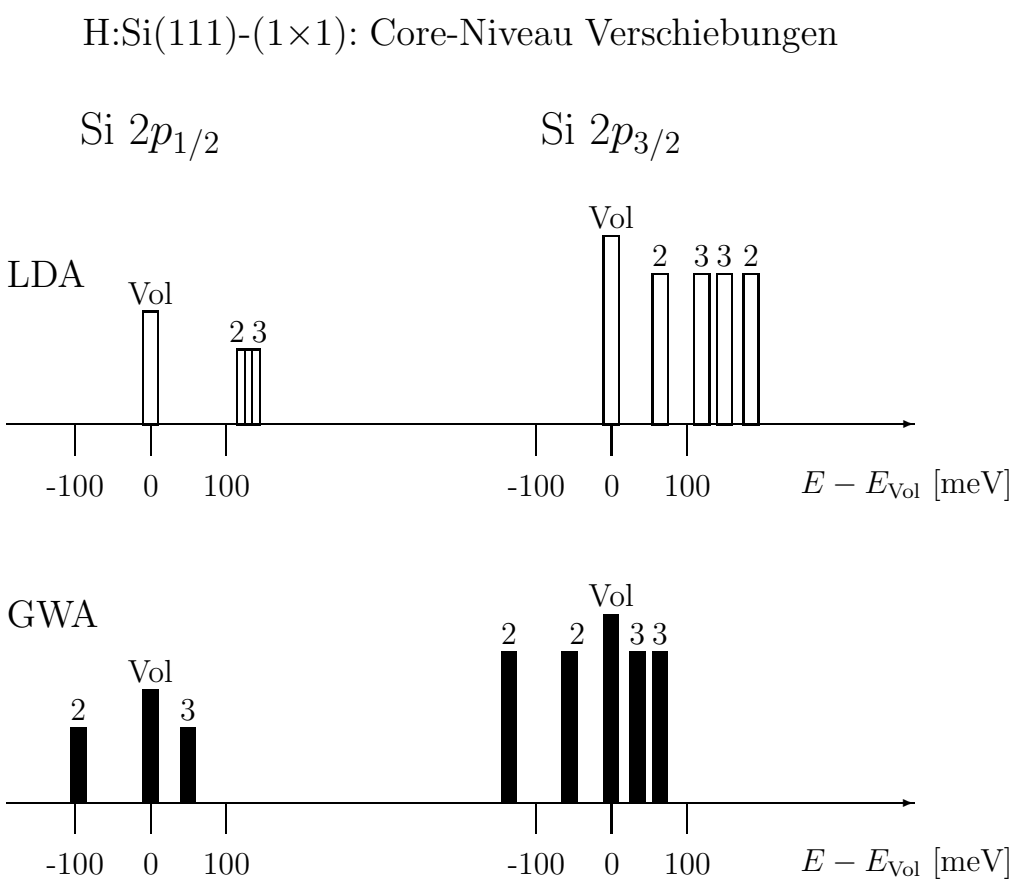
$$\Delta E^{(i)}_{3/2,\pm 1/2} = 1/3\Delta E^{(i)}_{1,\pm 1} + 2/3\Delta E^{(i)}_{1,0} \quad (15.7)$$

$$\Delta E^{(i)}_{1/2,\pm 1/2} = 2/3\Delta E^{(i)}_{1,\pm 1} + 1/3\Delta E^{(i)}_{1,0} \quad . \quad (15.8)$$

Der im freien Atom und im Volumen vierfach entartete Zustand Si $2p_{3/2}$ spaltet an der Oberfläche also in zwei Niveaus auf und wird außerdem noch verschoben.

Wir stellen in Abb. 15.6 die Ergebnisse dar, die wir für die Core-Niveau Verschiebungen der beiden oberflächennächsten Si Atome erhalten. Im Rahmen der LDA ergeben sich positive Verschiebungen, d.h. die Energie der Si $2p$ Zustände wird an der Oberfläche angehoben; die Bindungsenergie wird reduziert. Darin unterscheidet sich die Oberfläche deutlich von der As:Si(111)-(1×1) Oberfläche, an der die Bindungsenergie verstärkt worden war. Dieser Unterschied war zunächst nicht zu vermuten, denn die Pauling-Elektronegativitäten von H und As sind gleich groß. Im Rahmen der LDA lassen sich die Verschiebungen wieder als rein elektrostatischer Effekt identifizieren. Das Austausch-Korrelations Potential trägt nur mit 10 meV zu den Verschiebungen gegenüber dem Volumen bei. Die LDA-Daten für die Verschiebungen weichen von den gemessenen Verschiebungen [Bla94b, Dum96, Hor96] um ca. 250 meV ab.

Durch die GWA Quasiteilchen-Korrekturen tritt eine deutliche Änderung ein. An den Oberflächenatomen erfahren die Si $2p$ Zustände Quasiteilchen-Korrekturen zur LDA, die um –220 meV (an Si 2) bzw. –85 meV (an Si 3) stärker ausfallen als die Quasiteilchen-Korrekturen im Volumen. Um den gleichen Wert ändern sich auch die Core-Niveau Verschiebungen gegenüber denen, die sich aus der LDA-Bandstruktur ergeben. Diese Effekte sind wesentlich größer als an der As:Si(111)-(1×1) Oberfläche, an der die zusätzliche Niveauverschiebung durch die Quasiteilchen-Effekte



	Si 2 (Oberfläche)			Si 3 (Sublayer)		
[meV]	LDA	GWA	Exp.	LDA	GWA	Exp.
Si $2p_{3/2}$	+65	–135	–150[a], –235[b]	+120	+35	+90[a], –145[b]
	+185	–55	–80[a], –235[b]	+150	+65	+90[a], –145[b]
Si $2p_{1/2}$	+125	–95	–120[a], –235[b]	+135	+50	+90[a], –145[b]

[a] Ref. [Bla94b, Dum96] [b] Ref. [Hor96]

Abbildung 15.6: *Core-Niveau Verschiebungen an der H:Si(111)-(1×1) Oberfläche (siehe Text).*

nur 50 meV betragen hatte. Offensichtlich sind die Umgebung der Si Atome und die Abschirmung an der H-terminierten Oberfläche deutlich vom Volumen verschieden.

Die Quasiteilchen-Verschiebungen stimmen wesentlich besser mit dem Experiment überein als die LDA-Daten, liegen aber tendentiell ebenfalls über den Meßdaten. Die Messungen selbst ergeben allerdings kein vollständig einheitliches Bild

der Oberfläche. In den Messungen von Horn [Hor96] liegen die Core-Niveaus an der Oberfläche um mehr als 100 meV unter den in [Bla94b, Dum96] angegebenen. Die große Kristallfeld-Aufspaltung des Si $2p_{3/2}$ Niveaus am äußersten Si Atom in der letzteren Messung wird von Horn nicht beobachtet. Wir ermitteln hierfür einen Wert von 80 meV in sehr guter Übereinstimmung mit dem in [Bla94b, Dum96] gemessenen Wert von 70 meV. Abschließend vergleichen wir die mittleren Verschiebungen an den einzelnen Atomen (Si 2 und Si 3) miteinander. Sie unterscheiden sich in unserer Rechnung um 145 meV voneinander. Im Experiment werden hierfür Differenzen von 210 meV [Bla94b, Dum96] bzw. 90 meV [Hor96] beobachtet.

Unsere Ergebnisse sollten allerdings nicht überinterpretiert werden, denn die vorliegenden Core-Niveau Verschiebungen sind sehr klein und liegen fast schon in der Größenordnung unserer Rechengenauigkeit, die wir auf etwa 50 meV schätzen.

15.3 Die reine Si(001)-(2×1) Oberfläche

Zum Abschluß dieses Kapitels untersuchen wir die Core-Niveau Verschiebungen an der reinen Si(001)-(2×1) Oberfläche. Auch für diese Oberfläche hatten wir in Kap. 7 bereits ausführlich die Eigenschaften der Valenz- und Leitungszustände diskutiert. Das Core-Niveau Spektrum dieses Systems zeichnet sich durch zwei deutlich getrennte oberflächeninduzierte Emissionslinien fast gleicher Intensität aus [Him80]. Daraus wurde geschlossen, daß an der Oberfläche Si Atome in unterschiedlichen Umgebungen im Verhältnis 1:1 vorliegen müssen. Das Modell des asymmetrischen Dimers vermag diesen Befund zu erklären. Durch die Asymmetrie des Dimers kommt es zu einem Ladungstransport vom Down-Atom zum Up-Atom. Dadurch wird das elektrostatische Potential für die Elektronen am Up-Atom angehoben und am Down-Atom abgesenkt. Wir erwarten daher für das Si $2p$ Niveau am Up-Atom eine Anhebung und am Down-Atom eine Absenkung gegenüber dem Volumen.

Zur Klärung dieser Frage haben wir analog zu den adsorbat-bedeckten Si(111) Oberflächen diese Core-Niveaus berechnet. Wir konzentrieren uns dabei auf die beiden unterschiedlichen Dimeratome und auf ein Referenzatom in der Slabmitte. Wir stellen in Tabelle 15.1 die erzielten Ergebnisse dar. Die Verschiebungen, die wir in LDA erhalten, gehen wieder nahezu vollständig auf das elektrostatische Potential zurück. Das Austausch-Korrelations Potential beeinflußt die Verschiebungen nur in der Größenordnung 0.05 eV. Ein interessanter Aspekt betrifft die äußerst geringe Kristallfeld-Aufspaltung der Core-Zustände. Sie beträgt sowohl an den Down-

Si(001)-(2×1): Core-Niveau Verschiebungen

[eV]	LDA	GWA	Exp.
Down-Atom	–0.76	–0.59	–0.06
Up-Atom	0.51	1.06	0.48
Mittelwert	–0.13	0.24	0.21

Tabelle 15.1: *Core-Niveau Verschiebungen an der Si(001)-(2×1) Oberfläche (siehe Text). Die experimentellen Daten wurden von Landemark et al. bestimmt [Lan92].*

Atomen als auch an den Up-Atomen weniger als 0.05 eV. Das ist besonders für die Down-Atome ein überraschendes Ergebnis, denn sie weisen aufgrund der Asymmetrie des Dimers keine tetraedrische Konfiguration mehr auf.

Durch die Quasiteilchen-Korrekturen werden die Core-Niveau Verschiebungen relativ zum Volumen deutlich verändert. Da die fraglichen Atome sich an der freien Oberfläche direkt in der Oberflächenebene befinden, weichen die für die Core-Zustände relevanten Eigenschaften des Selbstenergie-Operators wesentlich stärker vom Volumen ab, als das bei den Si Sublayer-Atomen der adsorbat-bedeckten Oberflächen der Fall war. Durch die Quasiteilchen-Effekte wird die negative Verschiebung des Si $2p$ Niveaus am Down-Atom um 0.17 eV reduziert. Die positive Verschiebung am Up-Atom steigt gleichzeitig um +0.55 eV an. Die endgültigen Core-Niveau Verschiebungen von –0.59 eV und +1.06 eV weichen deutlich von den gemessenen Verschiebungen ab. Die Abweichungen von 0.5 eV sind wesentlich größer als für die oben diskutierten Si(111) Oberflächen, bei denen die verbleibenden Fehler in der Größenordnung 0.1 eV lagen.

Wir merken zunächst an, daß sich der *Mittelwert* der von uns berechneten Verschiebungen für das Down- und das Up-Atom in GWA zu 0.24 eV ergibt. Dieser Wert stimmt sehr gut mit dem Mittelwert der gemessenen Verschiebungen überein, der 0.21 eV beträgt. Die Abweichungen der Up- und Down-Verschiebungen von diesem Mittelwert werden von unserer GW-Rechnung um einen Faktor 3 gegenüber dem Experiment überschätzt. Als mögliche Ursache für die Diskrepanz kommen Unsicherheiten in der Oberflächengeometrie in Betracht. Wir diskutieren hier zwei Aspekte.

(i) Möglicherweise hat der Dimer im Experiment einen anderen Kippwinkel, als wir

auf der Basis der LDA-Geometrie-Optimierung vorausgesetzt haben. Es erscheint plausibel, daß mit einem kleineren Kippwinkel auch ein geringerer Ladungstransfer verbunden ist. Dies führt letztlich zu einer Angleichung der Core-Niveau Verschiebungen an den Up- und Down-Atomen. Im Grenzfall symmetrischer Dimere würden beide Dimeratome die gleiche Niveau-Verschiebung erfahren, die in erster Näherung durch den von uns bestimmten Mittelwert gegeben wäre.
(ii) Die Ladungsverteilung der Oberfläche ist durch positive Partialladungen an den Down-Atomen und negative Partialladungen an den Up-Atomen gekennzeichnet. Dadurch bilden sich Dipole aus. Das elektrostatische Potential an einem Oberflächenatom hängt entscheidend davon ab, in welcher Rekonstruktion die benachbarten Dipole vorliegen. In unser Rechnung sind wir von einer (2×1) Ordnung ausgegangen. Einfache Überlegungen ergeben, daß durch einen Übergang von der (2×1) zur $c(4\times2)$ oder $p(2\times2)$ Geometrie die Struktur des Potentials derart verändert wird, daß die Core-Niveaus am Down-Atom angehoben und am Up-Atom abgesenkt werden. Landemark et al. haben bei der Messung des Core-Niveau Spektrums festgestellt, daß sich für die $c(4\times2)$-geordnete Oberfläche bei tiefer Temperatur das gleiche Spektrum ergibt wie für die nominale (2×1)-geordnete Oberfläche bei Raumtemperatur (RT) [Lan92]. Die Autoren schließen daraus, daß die Oberfläche auch bei RT eine lokale $c(4\times2)$ Geometrie aufweist (vgl. die Diskussion in Kap. 7). Pehlke und Scheffler [Peh93] haben Core-Niveau Verschiebungen für eine $p(2\times2)$ geordnete Oberfläche bestimmt, deren Dipolverteilung einer $c(4\times2)$ Oberfläche näher kommt als die der (2×1) Oberfläche. Die Autoren erhalten tatsächlich an den Down-Atomen ein höheres und an den Up-Atomen ein niedrigeres Core-Niveau als in unserer Rechnung für die (2×1) Oberfläche.

Wir haben in diesem Kapitel die Core-Niveau Verschiebungen an drei Oberflächen bestimmt. Die ersten beiden Systeme (As:Si(111)-(1×1) und H:Si(111)-(1×1)) zeichnen sich durch eine stabile Geometrie aus. Wir hatten in Kapitel 9 die Valenzband- und Leitungsbandstruktur beider Systeme ermittelt und dabei eine sehr gute Übereinstimmung mit dem Experiment erzielt. Daraus läßt sich schließen, daß die zugrunde liegende, in LDA bestimmte Geometrie ein sehr realistisches Bild der jeweiligen Oberfläche zeichnet. Dieser Schluß wird nun durch die gute Übereinstimmung der berechneten Core-Niveau Verschiebungen mit gemessenen Daten bestätigt. Gleichzeitig untermauert dies die Aussagekraft unseres Zugangs zur direkten Bestimmung der Quasiteilchen-Energien der Core-Zustände.

Im Falle der Si(001)-(2×1) Oberfläche liegt eine andere Situation vor. Hier be-

sitzt die Oberflächengeometrie eine kompliziertere Struktur, die durch höhere Rekonstruktionen und dynamische Vorgänge bei Raumtemperatur gekennzeichnet ist. In diesem Falle weisen die von uns in GWA bestimmten Core-Niveaus Abweichungen von 0.5 eV zum Experiment auf. Interessanterweise ließen sich die Quasiteilchen-Energien der Dangling-Bond Zustände weitaus zuverlässiger berechnen (vgl. Kap. 7). Wir führen das darauf zurück, daß die Dangling Bonds zu einem erheblichen Teil in das Substrat hineinreichen, während die Core-Zustände zu 100 % an den Dimeratomen lokalisiert sind. Die Core-Zustände reagieren daher auf Modifikationen der Eigenschaften der Dimeratome viel stärker als die Dangling-Bond Zustände. Letztlich stellen die Core-Zustände der Oberflächenatome ein sehr empfindliches Werkzeug dar, um durch Vergleich zwischen Theorie und Experiment direkte Information über die geometrische Struktur der Oberfläche zu erhalten.

Kapitel 16

Zusammenfassung

Gegenstand dieser Arbeit war die Bestimmung von elektronischen Quasiteilchen-Bandstrukturen im Rahmen eines ab-initio Zugangs. Die Austausch-Korrelations Beiträge zu den Bandstruktur-Energien wurden mittels der GW-Näherung für den Selbstenergie-Operator erfaßt. Die komplizierte Orts- und Zeitabhängigkeit dieses Operators erfordert die Einführung geeigneter Darstellungsmethoden. Die Zeit- bzw. Frequenzabhängigkeit wurde mit Hilfe von Plasmon-Pol-Modellen beschrieben. Die Ortsabhängigkeiten wurden in dieser Arbeit erstmalig mit lokalisierten Gauß-Orbitalen erfaßt, die konsequent in allen Rechenschritten zur Anwendung gelangten. Dieser Zugang erweist sich besonders bei solchen Systemen als nützlich, die lokalisierte Zustände aufweisen. Bei Oberflächensystemen läßt sich eine beachtliche Verkürzung der Rechenzeiten im Vergleich zu herkömmlichen Ebene-Wellen Rechnungen erzielen. Bei der Untersuchung von Semicore- und Core-Zuständen stellt die Verwendung lokalisierter Ansatzfunktionen eine Grundvoraussetzung zur Durchführbarkeit der GWA dar. Die konkrete Umsetzung der Untersuchungen geschah auf der Grundlage vorangestellter Dichtefunktionaltheorie-Rechnungen im Rahmen der lokalen-Dichte-Approximation (LDA). Es wurden Quasiteilchen-Korrekturen zu LDA-Bandstrukturen bestimmt.

Wir haben für eine Vielzahl unterschiedlicher Oberflächen von Si, Ge und SiC die Quasiteilchen-Bandstrukturen der Oberflächenzustände untersucht und miteinander verglichen. Die elektronische Struktur der Dimer-terminierten reinen Oberflächen Si(001)-(2×1), Ge(001)-(2×1) und β-SiC(001)-(2×1) (C-terminiert) wird von freien Bindungen (dangling bonds) dominiert, deren Bandstrukturenergien im Bereich des fundamentalen Gaps liegen. Die Quasiteilchen-Bandstrukturen dieser Oberflächen

sind halbleitend. Durch die Adsorption von As auf Si(111) bzw. S auf Ge(001) werden die Oberflächen chemisch passiviert und nehmen eine stabile (1×1) Rekonstruktion an. Die Zustände der doppelt besetzten freien Bindungen der Adatome weisen ebenfalls Energien im Gapbereich auf. Auch für diese Oberflächen ergeben sich halbleitende Quasiteilchen-Bandstrukturen. Falls hingegen Wasserstoff adsorbiert wird, so liegen an der Oberfläche keine freien Bindungen mehr vor. Die Energien der chemischen Bindungen zwischen den H-Adatomen und dem Substrat liegen einige eV unterhalb des Valenzbandmaximums, und der Gapbereich ist frei von Oberflächenzuständen.

Bis auf wenige Ausnahmen stimmen unsere berechneten GW Quasiteilchen-Energien sehr gut mit verfügbaren experimentellen Daten überein. Die GWA erweist sich als zuverlässiges Werkzeug zur Bestimmung von Bandstrukturenergien lokalisierter Oberflächenzustände. Unser Zugang mittels lokalisierter Gauß-Orbital Basisfunktionen erlaubt zudem eine quantitativ verläßliche Auswertung der GWA.

Die Verwendung lokalisierter Ansatzfunktionen ermöglichte uns außerdem eine äußerst effiziente Behandlung stark lokalisierter Semicore- und Core-Zustände im Rahmen der GWA. Zunächst haben wir die Bedeutung von Semicore-Zuständen für die Quasiteilchen-Bandstruktur untersucht. Als wichtiges Ergebnis stellen wir fest, daß im Rahmen unseres Pseudopotentialzugangs alle Semicore-Zustände einer Schale in gleicher Weise behandelt werden müssen, um zu einer verläßlichen Berechnung der Bandstrukturen zu gelangen. Die Unterteilung einer Schale in Core- und Valenzzustände führt zu unkontrollierbaren systematischen Abweichungen in den Quasiteilchen-Energien. Die Einbeziehung sämtlicher Zustände der Semicore-Schale in das Valenzelektronensystem führt hingegen auf quantitativ richtige Bandstrukturen und Gapenergien. Die Auswirkungen der Semicore-Zustände auf die Quasiteilchen-Bandstrukturen weisen systematische Trends innerhalb des Periodensystems auf. Innerhalb einer Periode nimmt der Einfluß von links nach rechts ab [z.B. Zn → Ga → Ge], innerhalb einer Gruppe nimmt er von oben nach unten zu [z.B. Si → Ge → Sn].

Die Berechnung von GW Quasiteilchen-Bindungsenergien von Core-Zuständen erfordert eine Modifikation der Auswertung der GWA in Richtung auf einen selbstkonsistenten Selbstenergie-Operators. Dieser muß das selbstkonsistente Quasiteilchen-Spektrum beinhalten und sollte Satellitenstrukturen berücksichtigen. Die damit bestimmten Bindungsenergien reproduzieren die experimentellen Daten. Ab-

schließend haben wir die chemische Verschiebung von Core-Niveaus an Oberflächen im Rahmen der GWA diskutiert. Wegen der vollständigen Lokalisierung der Core-Zustände an den einzelnen Oberflächenatomen ermöglicht eine solche Untersuchung genaue Aussagen über die geometrische Struktur der Oberfläche.

Die GW Näherung erweist sich als eine universelle Methode zur zuverlässigen Bestimmung von Quasiteilchen-Energien unterschiedlicher elektronischer Zustände verschiedenster Festkörpersysteme. Sie läßt sich somit in die Standard-Werkzeuge zur Untersuchung kondensierter Materie einordnen.

Anhang A

Darstellung von Zweipunkt-Funktion

In diesem Abschnitt soll dargestellt werden, wie mit einem ausreichend flexiblen Satz von Basisfunktionen gitterperiodische Einpunkt- und Zweipunkt-Funktionen darstellen werden können.

Gegeben sei ein Basissatz $\{\chi_\beta(\mathbf{q},\mathbf{r})\}$. Die Funktionen sollen das Bloch-Theorem erfüllen, d.h. der Ausdruck $e^{-i\mathbf{qr}}\chi_\beta(\mathbf{q},\mathbf{r})$ stelle eine gitterperiodische Funktion dar. Der Index β indiziert die verschiedenen Funktionen zu einem gegebenen $\mathbf{q}$ aus der ersten Brillouinzone. Falls die Indexmenge β endlich ist (das ist bei konkreten Basissätzen immer der Fall), so liegt keine Vollständigkeit vor. Die Funktionen χ_β sind im allgemeinen nicht orthogonal.

Gegeben sei ferner eine (komplexwertige) Zweipunktfunktion $f(\mathbf{r},\mathbf{r}')$, die gitterperiodisch, also invariant unter einer Gittertranslation $\mathbf{R}$ sein soll:

$$f(\mathbf{r}+\mathbf{R},\mathbf{r}'+\mathbf{R}) = f(\mathbf{r},\mathbf{r}') \quad . \tag{A.1}$$

Die Darstellung einer solchen Funktion kann mit Hilfe des Basissatzes $\{\chi_\beta(\mathbf{q},\mathbf{r})\}$ erfolgen:

$$f(\mathbf{r},\mathbf{r}') = \sum_{\mathbf{q}} \sum_{\beta\beta'} \chi_\beta(\mathbf{q},\mathbf{r}) \langle f \rangle_{\beta\beta'}(\mathbf{q}) \chi^*_{\beta'}(\mathbf{q},\mathbf{r}') \quad . \tag{A.2}$$

Dies entspricht einer Transformation in beiden Ortsvariablen $\mathbf{r}$ und $\mathbf{r}'$. Die Gitterperiodizität der Zweipunkt-Funktion reduziert die Brillouinzonen-Integration auf nur eine Summe über $\mathbf{q}$.

Auf Gl. (A.2) wende man die Operationen $\int \chi^*_{\beta''}(\mathbf{q},\mathbf{r}) d^3r$ (von links) und

$\int \chi_{\beta'''}(\mathbf{q},\mathbf{r}')d^3r'$ (von rechts) an:

$$\int \chi^*_{\beta''}(\mathbf{q},\mathbf{r})f(\mathbf{r},\mathbf{r}')\chi_{\beta'''}(\mathbf{q},\mathbf{r}')d^3rd^3r' \;=\; \sum_{\beta\beta'}\int \chi^*_{\beta''}(\mathbf{q},\mathbf{r})\chi_{\beta}(\mathbf{q},\mathbf{r})d^3r\langle f\rangle_{\beta\beta'}(\mathbf{q})\int \chi^*_{\beta'}(\mathbf{q},\mathbf{r}')\chi_{\beta'''}(\mathbf{q},\mathbf{r}')d^3r'$$

$$\text{bzw.}\quad [f]_{\beta''\beta'''}(\mathbf{q}) \;=\; \sum_{\beta\beta'} S_{\beta''\beta}(\mathbf{q})\langle f\rangle_{\beta\beta'}(\mathbf{q})S_{\beta'\beta'''}(\mathbf{q}) \quad . \tag{A.3}$$

Die Brillouinzonen-Integration $\sum_{\mathbf{q}}$ entfällt infolge des Bloch-Theorems. Die linke Seite wurde durch den Ausdruck $[f]$ abgekürzt. Auf der rechten Seite wurde die Überlappmatrix S eingeführt:

$$S_{\beta\beta}(\mathbf{q}) = \int \chi^*_{\beta}(\mathbf{q},\mathbf{r})\chi_{\beta'}(\mathbf{q},\mathbf{r})d^3r \quad . \tag{A.4}$$

Diese Matrixgleichung kann umgeformt werden:

$$\text{aus}\quad [f] \;=\; S\langle f\rangle S \tag{A.5}$$

$$\text{folgt}\quad \langle f\rangle \;=\; S^{-1}[f]S^{-1} \quad . \tag{A.6}$$

Diese beiden Gleichungen stellen den Zusammenhang zwischen den Matrizen $[f]$ und $\langle f\rangle$ her. Formal setzt die Herleitung die exakte Gleichheit in Gl. (A.2) voraus. Das ist nur bei vollständigen Basissätzen erfüllt. Man erhält dieselben Beziehungen, wenn für einen unvollständigen Basissatz die mittlere quadratische Abweichung zwischen $f(\mathbf{r},\mathbf{r}')$ und $\sum \chi\langle f\rangle\chi^*$ minimiert werden soll:

$$\min_{\langle f\rangle_{\beta\beta'}} \int\int \left| f(\mathbf{r},\mathbf{r}') - \sum_{\mathbf{q}}\sum_{\beta\beta'} \chi_{\beta}(\mathbf{q},\mathbf{r})\langle f\rangle_{\beta\beta'}(\mathbf{q})\chi^*_{\beta'}(\mathbf{q},\mathbf{r}')\right|^2 d^3rd^3r' \quad .$$

Im folgenden geben wir einige Beziehungen an, die sich beim Rechnen mit den Matrizen S, $\langle f\rangle$ und $[f]$ als hilfreich erweisen. Wir notieren sie in Matrixschreibweise und verzichten weitgehend auf das Argument $\mathbf{q}$.

(1) Sei $f(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r},\mathbf{r}')$ die Delta-Distribution. Für ihre Matrixdarstellung gilt:

$$[\delta]_{\beta\beta'}(\mathbf{q}) \;=\; \int \chi^*_{\beta}(\mathbf{q},\mathbf{r})\delta(\mathbf{r},\mathbf{r}')\chi_{\beta'}(\mathbf{q},\mathbf{r})d^3rd^3r'$$

$$\;=\; S_{\beta\beta'}(\mathbf{q}) \quad . \tag{A.7}$$

Die Umrechnung (A.6) zur Matrix $\langle\delta\rangle$ führt auf

$$\langle\delta\rangle = S^{-1}SS^{-1} = S^{-1} \quad . \tag{A.8}$$

(2) Faltung: Sei f eine Faltung von g mit h: $f = g \circ h$; in expliziter Form:

$$f(\mathbf{r}, \mathbf{r}') = \int g(\mathbf{r}, \mathbf{r}'')h(\mathbf{r}'', \mathbf{r}')d^3r'' \quad . \tag{A.9}$$

Für die Matrixdarstellungen erhält man die Beziehungen

$$[f] \;=\; [g \circ h] = [g]S^{-1}[h] \tag{A.10}$$

$$\text{und} \quad \langle f \rangle \;=\; \langle g \circ h \rangle = \langle g \rangle S \langle h \rangle \quad . \tag{A.11}$$

(3) Inversion: eine Funktion f ist über die folgende Faltung mit ihrer Inversen f^{-1} verknüpft:

$$f \circ f^{-1} = f^{-1} \circ f = \delta \tag{A.12}$$

bzw.

$$\int f(\mathbf{r}, \mathbf{r}'')f^{-1}(\mathbf{r}'', \mathbf{r}')d^3r'' = \delta(\mathbf{r}, \mathbf{r}') \quad . \tag{A.13}$$

Mit Hilfe der obigen Regeln für Faltungen ergibt sich daraus:

$$[f]S^{-1}[f^{-1}] = [\delta] = S \tag{A.14}$$

bzw.

$$[f^{-1}] = S[f]^{-1}S = \left(S^{-1}[f]S^{-1}\right)^{-1} = \langle f \rangle^{-1} \tag{A.15}$$

und analog dazu

$$\langle f^{-1} \rangle = [f]^{-1} \quad . \tag{A.16}$$

(4) Umrechnung in Fouriertransformation: In vielen Fällen ist die Funktion f als Fouriertransformation $f_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ gegeben und soll in die Gauß-Darstellung umgerechnet werden oder umgekehrt. Das geschieht mittels der (analytisch bekannten) Fouriertransformation $K_\beta^{\mathbf{G}}(\mathbf{q})$ der Gauß-Orbitale $\chi_\beta(\mathbf{q}, \mathbf{r})$ (vgl. Anhang B):

$$[f]_{\beta\beta'}(\mathbf{q}) \;=\; \sum_{\mathbf{G},\mathbf{G}'} \left(K_\beta^{\mathbf{G}}(\mathbf{q})\right)^* f_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) K_{\beta'}^{\mathbf{G}'}(\mathbf{q}) \tag{A.17}$$

$$\text{mit} \quad K_\beta^{\mathbf{G}}(\mathbf{q}) \;=\; \int e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}} \chi_\beta(\mathbf{q}, \mathbf{r}) d^3r \quad . \tag{A.18}$$

für $f(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r}, \mathbf{r}')$ erhält man die Beziehung

$$S_{\beta\beta'}(\mathbf{q}) = \sum_{\mathbf{G}} \left(K_\beta^{\mathbf{G}}(\mathbf{q})\right)^* K_{\beta'}^{\mathbf{G}}(\mathbf{q}) \quad . \tag{A.19}$$

Unter der Annahme, daß der Basissatz $\{\beta\}$ vollständig sei, kann man ferner zeigen, daß gilt:

$$\sum_{\beta\beta'} K_{\beta}^{\mathbf{G}}(\mathbf{q}) S_{\beta\beta'}^{-1}(\mathbf{q}) \left(K_{\beta'}^{\mathbf{G}'}(\mathbf{q})\right)^* = \delta_{\mathbf{G},\mathbf{G}'} \quad . \tag{A.20}$$

Da die Gauß-Orbitale keine vollständige Basis bilden, gilt diese Beziehung nicht exakt, sondern nur im Rahmen einer gewissen Ungenauigkeit.

In Abschnitt 12.1 haben wir erläutert, wie im Rahmen der LDA Gauß-Orbitale zur Darstellung von Ladungsdichte und Potentialen verwendet werden können. Das beruht allgemein darauf, daß eine gitterperiodische Einpunkt-Funktion $f(\mathbf{r})$ durch die Basisfunktionen $\chi_\gamma(\mathbf{q}=\mathbf{0},\mathbf{r})$ dargestellt werden kann:

$$f(\mathbf{r}) = \sum_{\gamma} \langle f \rangle_\gamma \chi_\gamma(\mathbf{0},\mathbf{r}) \quad . \tag{A.21}$$

In gleicher Weise wie bei den Zweipunkt-Funktionen zeigt man, daß gilt:

$$\langle f \rangle \;=\; S^{-1}(\mathbf{0})[f] \tag{A.22}$$

$$\text{mit} \quad [f]_\gamma \;=\; \int \chi_\gamma^*(\mathbf{0},\mathbf{r}) f(\mathbf{r}) d^3r \quad . \tag{A.23}$$

Falls die Funktion $f(\mathbf{r})$ reell ist, so sind unter Verwendung der Gauß-Orbitale auch die Matrizen $[f]$ und $\langle f \rangle$ reellwertig. Auch hier besteht die Möglichkeit, aus einer gegebenen PW-Darstellung $\{f_{\mathbf{G}}\}$ in die Gauß-Darstellung zu transformieren:

$$[f]_\gamma = \sum_{\mathbf{G}} \left(K_\gamma^{\mathbf{G}}(\mathbf{0})\right)^* f_{\mathbf{G}} \quad . \tag{A.24}$$

Anhang B

Die Ansatzfunktionen

Die Ansatzfunktionen $\chi_\alpha(\mathbf{q},\mathbf{r})$ sind durch Linearkombinationen von Gauß-Orbitalen gegeben, die an den Atompositionen $\boldsymbol{\tau}_\alpha$ zentriert sind:

$$\chi_\alpha(\mathbf{q},\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{q}(\mathbf{R}+\boldsymbol{\tau}_\alpha)} \phi_\alpha(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_\alpha) \tag{B.1}$$

Sowohl die GO-LDA Basis $\{\alpha\}$ als auch die GO-GWA Basis $\{\beta\}$ sind so aufgebaut. Wir verwenden in unseren Rechnungen Gauß-Orbitale mit den Drehimpulsen l=0—4 (s, p, d, f und g). In der LDA finden nur s-, p- und d-artige Funktionen Verwendung (maximaler Drehimpuls $l_{\max}$=2). Das häufig als s^* bezeichnete Orbital besitzt ebenfalls s-Symmetrie (l=0). Diese zehn Orbitale sind in Tabelle B.1 verzeichnet. In dieser Form sind die Gauß-Orbitale normiert.

Die Fouriertransformation einer Basisfunktionen χ_α ist durch

$$\begin{aligned} K_\alpha^{\mathbf{G}}(\mathbf{q}) &= \frac{1}{V} \int_V e^{-i\mathbf{G}\mathbf{r}} \sum_{\mathbf{R}} e^{-i\mathbf{q}(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_\alpha)} \phi_\alpha(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_\alpha) d^3r \\ &= \frac{1}{V} e^{-i\mathbf{G}\boldsymbol{\tau}_\alpha} \phi_\alpha(\mathbf{q}+\mathbf{G}) \end{aligned} \tag{B.2}$$

gegeben. Die hierbei auftretenden Fouriertransformierten $\phi_\alpha(\mathbf{k})$ der einzelnen Gauß-Orbitale sind für die s-, p-, d- und s^*-Orbitale ebenfalls in Tabelle B.1 verzeichnet.

Zusätzlich zu diesen zehn Funktionen sind zur Darstellung der GWA-Zweipunktfunktionen Basisfunktionen mit höheren Drehimpulsen notwendig, und zwar mit Drehimpuls l=3 (f-Funktionen) und l=4 (g-Funktionen) – letztere besonders bei Systemen mit d-Elektronen. Ein bezüglich der Drehoperationen vollständiger Basissatz zum Drehimpuls l beinhaltet $2l$+1 Funktionen. Wir verwenden in dieser Arbeit jedoch mehr als sieben bzw. neun Funktionen, um diese Drehimpulse l=3 bzw. l=4

L	Nr.	ijk	Orbital	$\phi(\mathbf{r})$ $[\times(2\alpha/\pi)^{3/4}e^{-\alpha\mathbf{r}^2}]$	$\phi(\mathbf{k})$ $[\times(2\pi/\alpha)^{3/4}e^{-\mathbf{k}^2/4\alpha}]$
0	1	000	s	1	1
1	2	100	p_x	$2\sqrt{\alpha}x$	$-ik_x/\sqrt{\alpha}$
	3	010	p_y	$2\sqrt{\alpha}y$	$-ik_y/\sqrt{\alpha}$
	4	001	p_z	$2\sqrt{\alpha}z$	$-ik_z/\sqrt{\alpha}$
2	5	101	d_{xz}	$4\alpha xz$	$-k_xk_z/\alpha$
	6	011	d_{yz}	$4\alpha yz$	$-k_yk_z/\alpha$
	7	110	d_{xy}	$4\alpha xz$	$-k_xk_y/\alpha$
	8		$d_{3z^2-r^2}$	$2\alpha(3z^2-r^2)/\sqrt{3}$	$(\mathbf{k}^2-3k_z^2)/(2\sqrt{3}\alpha)$
	9		$d_{x^2-y^2}$	$2\alpha(x^2-y^2)$	$(k_y^2-k_x^2)/(2\alpha)$
	10		s^*	$4\alpha\mathbf{r}^2/\sqrt{15}$	$(6-\mathbf{k}^2/\alpha)/\sqrt{15}$

Tabelle B.1: *Liste der zehn Gauß-Orbitale mit den Drehimpulsen l=0,1,2. Rechte Spalte: Fouriertransformation.*

abzudecken, und zwar aus folgendem Grunde. Die Ansatzfunktionen sind in unseren Rechnungen durch *kartesische* Gauß-Orbitale gegeben, die sich allgemein in der Form

$$\phi_{ijk}(\mathbf{r}) = A_{ijk}x^iy^jz^ke^{-\alpha\mathbf{r}^2} \tag{B.3}$$

darstellen lassen. Alle Funktionen mit $i+j+k=L$ bilden dabei einen Basissatz, der einerseits alle Winkelfunktionen zum Drehimpuls $l=L$ beinhaltet, gleichzeitig aber auch noch Anteile geringeren Drehimpulses enthält. Für $L=0$ ergibt sich die s-Funktion ϕ_1, für $L=1$ ergeben sich die p-Funktionen ϕ_2 bis ϕ_4. Die sechs Funktionen zu $L=2$ (nämlich x^2, y^2, z^2, xy, xz und yz) lassen sich zu den fünf d-Funktionen ϕ_5 bis ϕ_9 linear kombinieren (vgl. Tabelle B.1), beinhalten darüber hinaus aber auch noch die s^*-Funktion ϕ_{10}.

Die zehn Funktionen zu $L=3$ (x^3, ..., x^2y, ..., xyz etc.) enthalten einerseits die sieben f-Orbitale, des weiteren aber auch noch p-artige Anteile. Man könnte aus den kartesischen Funktionen durch Linearkombination die sieben f-Orbitale extrahieren und den Rest unterdrücken. Da wir aber viele Rechenschritte mit den kartesischen Funktionen besonders schnell ausführen können (siehe nächsten Abschnitt), und weil die weiteren Anteile die Flexibilität des Basissatzes weiter steigern, verzichten wir auf diese Zerlegung. Zur numerischen Darstellung der fraglichen Funktionen sind die zehn Basisfunktionen zu $L=3$ besser geeignet als nur die sieben darin enthaltenen f-Funktionen. Das gleiche gilt für die 15 Funktionen zu $L=4$ und die darin enthaltenen neun g-Orbitale. Aus ähnlichen Gründen verzichten wir auch auf die individuelle

L	Nr.	ijk	$\phi(\mathbf{r})$ $[\times(2\alpha/\pi)^{3/4}e^{-\alpha\mathbf{r}^2}]$	$\phi(\mathbf{k})$ $[\times 1/8\alpha^3(2\pi/\alpha)^{3/4}e^{-\mathbf{k}^2/4\alpha}]$
3	11	300	x^3	$ik_x(k_x^2-6\alpha)$
	12	030	y^3	$ik_y(k_y^2-6\alpha)$
	13	003	z^3	$ik_z(k_z^2-6\alpha)$
	14	210	x^2y	$i(k_x^2-2\alpha)k_y$
	15	120	xy^2	$ik_x(k_y^2-2\alpha)$
	16	201	x^2z	$i(k_x^2-\alpha)k_z$
	17	102	xz^2	$ik_x(k_z^2-2\alpha)$
	18	021	y^2z	$i(k_y^2-\alpha)k_z$
	19	012	yz^2	$ik_y(k_z^2-2\alpha)$
	20	111	xyz	$ik_xk_yk_z$

Tabelle B.2: *Liste der zehn kartesischen Gauß-Orbitale zu L=3, die die f-Orbitale enthalten.*

L	Nr.	ijk	$\phi(\mathbf{r})$ $[\times(2\alpha/\pi)^{3/4}e^{-\alpha\mathbf{r}^2}]$	$\phi(\mathbf{k})$ $[\times 1/16\alpha^4(2\pi/\alpha)^{3/4}e^{-\mathbf{k}^2/4\alpha}]$
4	21	400	x^4	$12\alpha(\alpha-k_x^2)+k_x^4$
	22	040	y^4	$12\alpha(\alpha-k_y^2)+k_y^4$
	23	004	z^4	$12\alpha(\alpha-k_z^2)+k_z^4$
	24	310	x^3y	$(k_x^2-6\alpha)k_xk_y$
	25	130	xy^3	$k_x(k_y^2-6\alpha)k_y$
	26	301	x^3z	$(k_x^2-6\alpha)k_xk_z$
	27	103	xz^3	$k_x(k_z^2-6\alpha)k_z$
	28	031	y^3z	$(k_y^2-6\alpha)k_yk_z$
	29	013	yz^3	$k_y(k_z^2-6\alpha)k_z$
	30	202	x^2z^2	$(k_x^2-2\alpha)(k_z^2-2\alpha)$
	31	022	y^2z^2	$(k_y^2-2\alpha)(k_z^2-2\alpha)$
	32	220	x^2y^2	$(k_x^2-2\alpha)(k_y^2-2\alpha)$
	33	211	x^2yz	$(k_x^2-2\alpha)k_yk_z$
	34	121	xy^2z	$k_x(k_y^2-2\alpha)k_z$
	35	112	xyz^2	$k_xk_y(k_z^2-2\alpha)$

Tabelle B.3: *Liste der fünfzehn kartesischen Gauß-Orbitale zu L=4, die die g-Orbitale enthalten.*

Normierung jeder Funktion und setzen $A_{ijk}=(2\alpha/\pi)^{3/4}$ für L=3 und L=4. In Tabelle B.2 und B.3 geben wir diese 25 Funktionen und ihre Fouriertransformationen an.

Die verwendeten Abklingkonstanten der Gauß-Orbitale orientieren sich primär an den Eigenschaften der Atome, an denen die Funktionen zentriert sind. In Tabelle

Element	PP	α (LDA)	β (GWA)		
		$L\leq 2$ (s,p,d,s^*)	$L\leq 2$ (s,p,d,s^*)	L=3 (f)	L=4 (g)
C	4+	0.35, 1.7 0.25, 1.0, 2.86	0.35, 1.7	1.7	
Si	4+	0.15, 0.5 0.2, 0.6	0.15, 0.5 0.2, 0.6	0.5 0.6	
	12+	0.2, 0.5, 1.2, 2.7, 6, 14	0.2, 0.5, 1.32, 3.46, 9.12, 24	0.5 0.5	
Ge	4+	0.19, 0.5 0.16, 0.3, 0.6	0.19, 0.5	0.5	
	22+	0.16, 0.35, 0.77, 1.7, 3.7, 8.2, 18	0.18, 0.5, 2.2, 10	0.18, 0.5, 2.2, 10	0.18, 0.5, 2.2, 10
Sn	4+	0.14, 0.3, 0.7			
	22+	0.14, 0.3, 0.7, 2.5, 6, 14	0.16, 0.5, 2.2, 10	0.16, 0.5, 2.2, 10	0.16, 0.5, 2.2, 10
Cd	2+	0.18, 0.4, 0.8	0.2, 0.5	0.5	
	12,18,20+	0.18, 0.4, 0.8, 2, 4.5, 10	0.2, 0.5, 2.2, 10	0.2, 0.5, 2.2, 10	0.2, 0.5, 2.2, 10
Zn	2+	0.18, 0.39, 0.84	0.18, 0.39, 0.84	0.39, 0.84	
	20+	0.18, 0.39, 0.84, 1.8, 3.8, 8.4, 18	0.2, 0.5, 2.2, 10	0.2, 0.5, 2.2, 10	0.2, 0.5, 2.2, 10
Ga	3+	0.2, 0.5	0.2, 0.5	0.5	
	21+	0.2, 0.5, 2, 4.5, 9, 18	0.2, 0.5, 2.2, 10	0.2, 0.5, 2.2, 10	0.2, 0.5 2.2, 10
N	5+	0.2, 0.6, 1.4, 3	0.2, 0.6, 1.4, 3	0.6, 1.4, 3	
As	5+	0.2, 0.5 0.15, 0.35, 0.8	0.2, 0.5	0.5	
S	6+	0.18, 0.4, 0.9	0.18, 0.4, 0.9	0.4, 0.9	
H	1+	0.19, 1.16	0.19, 1.16		

Tabelle B.4: *Die in dieser Arbeit verwendeten Abklingkonstanten der Gauß-Orbitale.*

B.4 geben wir einen Überblick über die LDA- und GWA-Abklingkonstanten, die wir für die verschiedenen Elemente und Pseudopotentiale verwendet haben.

Anhang C

Rekursive Bestimmung von Matrixelementen

Die Berechnung von Matrixelementen verschiedener Operatoren in der Gauß-Basis stellt den größten Anteil an numerischem Aufwand innerhalb der GWA dar. Am schnellsten läßt sich das mit einer rekursiven Vorgehensweise realisieren [Oba86]. Wir stellen hier nur die wesentlichen Resultate vor. Für eine ausführliche Darstellung verweisen wir auf die Arbeit von Obara und Saiki [Oba86].
Gegeben seien kartesischen Gauß-Funktionen

$$\phi_{j,k,l}^{\alpha;\mathbf{A}}(\mathbf{r}) := (x - A_x)^j (y - A_y)^k (z - A_z)^l e^{-\alpha(\mathbf{r}-\mathbf{A})^2} \tag{C.1}$$

mit dem Zentrierungspunkt $\mathbf{A}$. Die Summe $L = j + k + l$ legt die Klasse dieser Funktion fest. Wir benötigen im wesentlichen die folgenden Integraltypen.
(1) Zweizentrenintegrale:
Ausgehend von dem Integral mit s-Funktionen

$$\begin{aligned}(\phi_{0,0,0}^{\alpha;\mathbf{A}}|\phi_{0,0,0}^{\beta;\mathbf{B}}) &= \int \phi_{0,0,0}^{\alpha;\mathbf{A}}(\mathbf{r})\phi_{0,0,0}^{\beta;\mathbf{B}}(\mathbf{r})d^3r \\ &= \left(\frac{\pi}{\alpha+\beta}\right)^{3/2} e^{-\frac{\alpha\beta}{\alpha+\beta}(\mathbf{A}-\mathbf{B})^2}\end{aligned} \tag{C.2}$$

bestimmt man alle weiteren Zweizentrenintegrale mit der Rekursionsformel

$$\begin{aligned}(\phi_{j+1,k,l}^{\alpha;\mathbf{A}}|\phi_{j',k',l'}^{\beta;\mathbf{B}}) &= \frac{\beta}{\alpha+\beta}(B_x - A_x)(\phi_{j,k,l}^{\alpha;\mathbf{A}}|\phi_{j',k',l'}^{\beta;\mathbf{B}}) \\ &+ \frac{1}{2(\alpha+\beta)}\left(j(\phi_{j-1,k,l}^{\alpha;\mathbf{A}}|\phi_{j',k',l'}^{\beta;\mathbf{B}}) + j'(\phi_{j,k,l}^{\alpha;\mathbf{A}}|\phi_{j'-1,k',l'}^{\beta;\mathbf{B}})\right) \quad .\end{aligned} \tag{C.3}$$

Hier wird letztlich von Integralen der Klassen $L+L'$ und $L+L'-1$ auf ein Integral der Klasse $L+L'+1$ geschlossen [L und L' sind die Klassen der beteiligten Basisfunktionen (jkl) und $(j'k'l')$]. Die Rekursionsvorschrift gilt in analoger Form auch für die anderen kartesischen Richtungen (also für k und l) und für Indexänderungen in der zweiten Funktion (j', k', l').

(2) Dreizentrenintegrale:

Auch hier kann man ausgehend vom Dreizentrenintegral dreier s-Funktionen

$$(\phi^{\alpha;\mathbf{A}}_{0,0,0}|\phi^{\beta;\mathbf{B}}_{0,0,0}|\phi^{\gamma;\mathbf{C}}_{0,0,0}) = \int \phi^{\alpha;\mathbf{A}}_{0,0,0}(\mathbf{r})\phi^{\beta;\mathbf{B}}_{0,0,0}(\mathbf{r})\phi^{\gamma;\mathbf{C}}_{0,0,0}(\mathbf{r})d^3r$$

$$= \left(\frac{\pi}{\alpha+\beta+\gamma}\right)^{3/2} e^{-\frac{\alpha\beta(\mathbf{A}-\mathbf{B})^2+\alpha\gamma(\mathbf{A}-\mathbf{C})^2+\beta\gamma(\mathbf{B}-\mathbf{C})^2}{\alpha+\beta+\gamma}} \tag{C.4}$$

eine Rekursionsvorschrift anwenden:

$$(\phi^{\alpha;\mathbf{A}}_{j+1,k,l}|\phi^{\beta;\mathbf{B}}_{j',k',l'}|\phi^{\gamma;\mathbf{C}}_{j'',k'',l''}) = \frac{\beta(B_x-A_x)+\gamma(C_x-A_x)}{\alpha+\beta+\gamma}(\phi^{\alpha;\mathbf{A}}_{j,k,l}|\phi^{\beta;\mathbf{B}}_{j',k',l'}|\phi^{\gamma;\mathbf{C}}_{j'',k'',l''})$$

$$+\frac{1}{2(\alpha+\beta+\gamma)}\Big[j(\phi^{\alpha;\mathbf{A}}_{j-1,k,l}|\phi^{\beta;\mathbf{B}}_{j',k',l'}|\phi^{\gamma;\mathbf{C}}_{j'',k'',l''})$$

$$+ j'(\phi^{\alpha;\mathbf{A}}_{j,k,l}|\phi^{\beta;\mathbf{B}}_{j'-1,k',l'}|\phi^{\gamma;\mathbf{C}}_{j'',k'',l''}) + j''(\phi^{\alpha;\mathbf{A}}_{j,k,l}|\phi^{\beta;\mathbf{B}}_{j',k',l'}|\phi^{\gamma;\mathbf{C}}_{j''-1,k'',l''})\Big] \quad . \tag{C.5}$$

Auch dieser Zusammenhang gilt analog für die anderen kartesischen Richtungen sowie für die mittlere und die rechte Funktion im Dreizentrenintegral.

(3) Zweizentrenintegrale mit ebenen Wellen:

An einigen Stellen sind Integrale der Form

$$(\phi^{\alpha;\mathbf{A}}_{j,k,l}|e^{i\mathbf{kr}}|\phi^{\beta;\mathbf{B}}_{j',k',l'}) = \int \phi^{\alpha;\mathbf{A}}_{j,k,l}(\mathbf{r})e^{i\mathbf{kr}}\phi^{\beta;\mathbf{B}}_{j',k',l'}(\mathbf{r})d^3r \tag{C.6}$$

zu berechnen. Hier ergibt sich das Integral für die s-Funktionen zu

$$(\phi^{\alpha;\mathbf{A}}_{0,0,0}|e^{i\mathbf{kr}}|\phi^{\beta;\mathbf{B}}_{0,0,0}) = (\phi^{\alpha;\mathbf{A}}_{0,0,0}|\phi^{\beta;\mathbf{B}}_{0,0,0})e^{\frac{i(\alpha\mathbf{kA}+\beta\mathbf{kB})-\mathbf{k}^2/4}{\alpha+\beta}} \quad , \tag{C.7}$$

und die Rekursionsvorschrift lautet

$$(\phi^{\alpha;\mathbf{A}}_{j+1,k,l}|e^{i\mathbf{kr}}|\phi^{\beta;\mathbf{B}}_{j',k',l'}) = \frac{\beta(B_x-A_x)+ik_x}{\alpha+\beta}(\phi^{\alpha;\mathbf{A}}_{j,k,l}|e^{i\mathbf{kr}}|\phi^{\beta;\mathbf{B}}_{j',k',l'})$$

$$+ \quad \frac{1}{2(\alpha+\beta)}\left(j(\phi^{\alpha;\mathbf{A}}_{j-1,k,l}|e^{i\mathbf{kr}}|\phi^{\beta;\mathbf{B}}_{j',k',l'}) + j'(\phi^{\alpha;\mathbf{A}}_{j,k,l}|e^{i\mathbf{kr}}|\phi^{\beta;\mathbf{B}}_{j'-1,k',l'}\right) \quad . \tag{C.8}$$

Anhang D

Dielektrische Matrix und Coulomb-Wechselwirkung

Ausgehend von der Polarisierbarkeit P wird die dielektrische Funktion ϵ als eine Faltung mit der Coulomb-Wechselwirkung berechnet: $\epsilon{=}1 - vP$. In dieser Form ist die dielektrische Matrix nicht symmetrisch (z. B. bzgl. $\mathbf{r}$ und $\mathbf{r}'$). In einer Matrixdarstellung, etwa in Fourierdarstellung, divergieren einige Matrixelemente im Grenzwert $\mathbf{q}{\to}\mathbf{0}$ wie $1/q$, da die Coulomb-Wechselwirkung wie $1/q^2$ divergiert. Man gelangt zu einer wesentlich stabileren und praktischeren Darstellung, wenn die dielektrische Matrix symmetrisiert wird ($\to \tilde{\epsilon}$). Wir stellen diesen Prozeß zunächst im Fourierraum dar:

$$\begin{aligned} \epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) &= \delta_{\mathbf{G},\mathbf{G}'} - \frac{4\pi e^2}{|\mathbf{q}+\mathbf{G}|^2} P_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) \\ \longrightarrow \tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) &:= \frac{|\mathbf{q}+\mathbf{G}|}{|\mathbf{q}+\mathbf{G}'|} \epsilon_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) \\ &= \delta_{\mathbf{G},\mathbf{G}'} - \frac{\sqrt{4\pi e^2}}{|\mathbf{q}+\mathbf{G}|} P_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) \frac{\sqrt{4\pi e^2}}{|\mathbf{q}+\mathbf{G}'|} \quad . \end{aligned} \tag{D.1}$$

Man gewinnt diese symmetrische Form also, indem die Polarisierbarkeit beidseitig mit einer Hilfsfunktion $\tilde{v}$ gefaltet wird, deren Fouriertransformation durch $\tilde{v}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \sqrt{4\pi e^2}/|\mathbf{q}+\mathbf{G}|\delta_{\mathbf{G},\mathbf{G}'}$ gegeben ist. Eine äquivalente Formulierung lautet: $\tilde{\epsilon} = \tilde{v}^{-1}\epsilon\tilde{v}$. In einer konkreten Basisdarstellung ist diese symmetrische dielektrische Funktion $\tilde{\epsilon}$ dann durch eine hermitesche Matrix gegeben, deren Elemente im Falle $\mathbf{q}{\to}\mathbf{0}$ endlich bleiben (siehe Anhang E).

Die Ortsdarstellung der Hilfsfunktion $\tilde{v}$ ergibt sich aus ihrer Fouriertransforma-

tion:

$$\begin{aligned} \tilde{v}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) &= \frac{\sqrt{4\pi e^2}}{|\mathbf{q}+\mathbf{G}|}\delta_{\mathbf{G},\mathbf{G}'} \\ \longrightarrow \tilde{v}(\mathbf{r},\mathbf{r}') &= \frac{e}{\pi^{3/2}\cdot|\mathbf{r}-\mathbf{r}'|^2} \quad . \end{aligned} \tag{D.2}$$

Diese Hilfsfunktion hängt mit der Coulomb-Wechselwirkung selber durch eine Faltung zusammen:

$$\begin{aligned} \sum_{\mathbf{G}''}\frac{\sqrt{4\pi e^2}}{|\mathbf{q}+\mathbf{G}|}\delta_{\mathbf{G},\mathbf{G}''}\frac{\sqrt{4\pi e^2}}{|\mathbf{q}+\mathbf{G}''|}\delta_{\mathbf{G}'',\mathbf{G}'} &= \frac{4\pi e^2}{|\mathbf{q}+\mathbf{G}|}\delta_{\mathbf{G},\mathbf{G}'} \\ \text{bzw.} \quad \tilde{v}\circ\tilde{v} &= v \\ \text{bzw.} \quad \int \tilde{v}(\mathbf{r},\mathbf{r}'')\tilde{v}(\mathbf{r}'',\mathbf{r}')d^3r'' &= v(\mathbf{r},\mathbf{r}') \quad . \end{aligned}$$

Die abgeschirmte Wechselwirkung W berechnet sich aus der ursprünglichen inversen dielektrischen Funktion ϵ^{-1} gemäß $W = \epsilon^{-1}v$, also durch Faltung mit der Coulomb-Wechselwirkung. Wenn hingegen die symmetrische inverse dielektrische Funktion $\tilde{\epsilon}^{-1}$ vorliegt, so muß beidseitig mit der Hilfsfunktion $\tilde{v}$ gefaltet werden, um die abgeschirmte Wechselwirkung zu erhalten:

$$W = \tilde{v}\tilde{\epsilon}^{-1}\tilde{v} = \tilde{v}(\tilde{v}^{-1}\epsilon\tilde{v})^{-1}\tilde{v} = \tilde{v}\tilde{v}^{-1}\epsilon^{-1}\tilde{v}\tilde{v} = \epsilon^{-1}v \quad . \tag{D.3}$$

Wir wenden uns nun der Berechnung von Matrixelementen der Hilfsfunktion $\tilde{v}$ in der von uns verwendeten Gauß-Basis $\{\beta\}$ zu. Obwohl sie im reellen Raum um eine Potenz schneller abfällt als die Coulomb-Wechselwirkung v, besitzt sie dennoch einen ausgeprägten langreichweitigen Charakter. Eine Berechnung der Matrixelemente

$$[\tilde{v}]_{\beta\beta'}(\mathbf{q}) = \sqrt{4\pi e^2}\int \chi^*_\beta(\mathbf{q},\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|^2}\chi_\beta(\mathbf{q},\mathbf{r}')d^3r d^3r' \tag{D.4}$$

führt auf eine Summe der Struktur $\sum_{\mathbf{R}} 1/\mathbf{R}^2$, deren konvergente Bestimmung äußerst unpraktikabel ist. Eine wesentlich einfachere Möglichkeit zur Berechnung der Matrixelemente bietet sich mittels der Fourierdarstellung an. Mit den Hilfsmitteln aus Anhang A gestaltet sich das wie folgt:

$$\begin{aligned} [\tilde{v}]_{\beta\beta'}(\mathbf{q}) &= \sum_{\mathbf{G},\mathbf{G}'}\left(K^{\mathbf{G}}_\beta(\mathbf{q})\right)^*\tilde{v}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})K^{\mathbf{G}'}_{\beta'}(\mathbf{q}) \\ &= \sum_{\mathbf{G}}\left(K^{\mathbf{G}}_\beta(\mathbf{q})\right)^*\frac{\sqrt{4\pi e^2}}{|\mathbf{q}+\mathbf{G}|}K^{\mathbf{G}}_{\beta'}(\mathbf{q}) \quad . \end{aligned} \tag{D.5}$$

Infolge der Isotropie der Hilfsfunktion $\tilde{v}$ reduziert sich die Summation zu einer einfachen Summe über $\mathbf{G}$, die sich numerisch äußerst schnell durchführen läßt.

Anhang E

Coulomb-Divergenz in der Brillouinzonen-Integration

Die Langreichweitigkeit der Coulomb-Wechselwirkung erschwert an einigen Stellen des GW Formalismus die konvergente Auswertung verschiedener Formeln erheblich. Konkret äußert sich das darin, daß im Falle $\mathbf{q}\to\mathbf{0}$ die Darstellungsmatrix der Coulomb-Wechselwirkung divergiert – sowohl in der Fourierdarstellung als auch in der von uns verwendeten Gauß-Orbital-Darstellung. Wir stellen in diesem Abschnitt zwei konkrete Probleme dar, die sich daraus ergeben, und erläutern Möglichkeiten, diese zu beheben.

Das erste Problem betrifft die konvergente Bestimmung der dielektrischen Matrix $[\tilde{\epsilon}](\mathbf{q})$ in der RPA im Falle $\mathbf{q}\to\mathbf{0}$, insbesondere die Ermittlung der makroskopischen dielektrischen Konstante. Es erweist sich als problematisch, diese Matrix für $\mathbf{q}\to\mathbf{0}$ gegen ihren richtigen Grenzwert $[\tilde{\epsilon}](\mathbf{0})$ konvergieren zu lassen.

Betrachten wir dieses Problem zunächst in der Fourierdarstellung. Hier gilt

$$\tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) = \delta_{\mathbf{G},\mathbf{G}'} - 4\pi e^2 \frac{1}{|\mathbf{q}+\mathbf{G}|} P_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) \frac{1}{|\mathbf{q}+\mathbf{G}'|} \quad . \tag{E.1}$$

Die Elemente der Polarisation verhalten sich für $\mathbf{q}\to\mathbf{0}$ folgendermaßen:

$$\begin{aligned}
&P_{\mathbf{0},\mathbf{0}}(\mathbf{q}) \sim q^2 \\
&P_{\mathbf{0},\mathbf{G}'}(\mathbf{q}) \sim q \\
&P_{\mathbf{G},\mathbf{0}}(\mathbf{q}) \sim q \\
&P_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) \sim \text{const.} \quad \text{für alle anderen } \mathbf{G},\, \mathbf{G}' \quad .
\end{aligned}$$

Die entsprechenden Elemente $\tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ konvergieren folglich für $\mathbf{q}\to\mathbf{0}$ gegen endliche Werte $\tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{0})$, ohne daß divergentes Verhalten auftritt. Die gleiche Stabilität

muß man auch von den Elementen $[\tilde{\epsilon}]_{\beta\beta'}(\mathbf{q})$ der Gauß-Matrix fordern. Wir stellen im folgenden dar, daß dennoch die Konvergenz im allgemeinen erst mit einigen zusätzlichen Schritten erreicht werden kann. Die Transformation in die Gauß-Basis kann formal wie folgt notiert werden (vgl. Anhang A):

$$[\tilde{\epsilon}]_{\beta\beta'}(\mathbf{q}) = \sum_{\mathbf{G},\mathbf{G}'} \left(K_\beta^{\mathbf{G}}(\mathbf{q})\right)^* \tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) K_{\beta'}^{\mathbf{G}'}(\mathbf{q}) \tag{E.2}$$

$$= S_{\beta\beta'}(\mathbf{q}) - \sum_{\mathbf{G},\mathbf{G}'} \left(K_\beta^{\mathbf{G}}(\mathbf{q})\right)^* \frac{1}{|\mathbf{q}+\mathbf{G}|} \underset{\uparrow}{P_{\mathbf{G},\mathbf{G}'}(\mathbf{q})} \frac{1}{\underset{\uparrow}{|\mathbf{q}+\mathbf{G}'|}} K_{\beta'}^{\mathbf{G}'}(\mathbf{q}) \quad . \tag{E.3}$$

Der Zusammenhang zwischen diesem Ausdruck und der Gleichung (3.17) wird durch die Summenregel (A.20) hergestellt. Diese *vollständige Eins* wird in Gleichung (E.3) an den mit Pfeilen gekennzeichneten Stellen eingefügt. Der Zusammenhang (A.20) gilt jedoch nicht exakt, denn die darin auftretende Gauß-Basis $\{\beta\}$ ist nicht vollständig. Die dadurch bedingten Abweichungen von der Gleichheit der linken mit der rechten Seite in (A.20) liegen in unseren Rechnungen in der Größenordnung 10^{-3}, sofern die Basis $\{\beta\}$ für das zu beschreibende System geeignet gewählt wird. Diese verbleibende Ungenauigkeit ist generell belanglos. Lediglich bei der Berechnung der dielektrischen Matrix für $\mathbf{q}\to\mathbf{0}$ gemäß Gleichung (E.3) treten sie in Erscheinung und führen zu einer gravierenden numerischen Instabilität. Betrachten wir beispielsweise die Summe über $\mathbf{G}$ in Gl. (E.3). In ihr werden – wegen der Abweichungen von der Vollständigkeit in Gl. (A.20) – Fourierelemente $P_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ mit $\mathbf{G}\neq\mathbf{0}$ (die für $\mathbf{q}\to\mathbf{0}$ nicht gegen Null konvergieren) fälschlicherweise mit dem ($\mathbf{G}=\mathbf{0}$)-Faktor $1/q$ verkoppelt. Auch wenn diese Verkopplung nur mit einem Kopplungsfaktor in der Größenordnung $< 10^{-3}$ ($\hat{=}$ Ungenauigkeit in Gl. A.20)) geschieht, so divergiert dennoch die gesamte Matrix $[\tilde{\epsilon}]_{\beta\beta'}(\mathbf{q})$ für $\mathbf{q}\to\mathbf{0}$ in nicht kontrollierbarer Weise. Dieses Verhalten tritt auch dann auf, wenn die dielektrische Matrix $[\tilde{\epsilon}]$ nach (3.17) direkt in der Gauß-Basis berechnet wird. Wir lösen dieses Problem folgendermaßen. Die Summe (E.2) wird im Falle $\mathbf{q}\to\mathbf{0}$ formal in zwei Anteile zerlegt:

$$\begin{aligned} [\tilde{\epsilon}]_{\beta\beta'}(\mathbf{q}) \;&=\; \sum_{\mathbf{G},\mathbf{G}'\neq\mathbf{0}} \left(K_\beta^{\mathbf{G}}(\mathbf{q})\right)^* \tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) K_{\beta'}^{\mathbf{G}'}(\mathbf{q}) \\ &+ \sum_{\substack{\mathbf{G},\,\mathbf{G}' \\ \mathbf{G}=\mathbf{0} \text{ oder } \mathbf{G}'=\mathbf{0}}} \left(K_\beta^{\mathbf{G}}(\mathbf{q})\right)^* \tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) K_{\beta'}^{\mathbf{G}'}(\mathbf{q}) \quad . \end{aligned} \tag{E.4}$$

Die erste dieser beiden Summen enthält keine kritischen Anteile und läßt sich di-

rekt in der Form $[\tilde{v}]S^{-1}[P]S^{-1}[\tilde{v}]$ auswerten: Man hat lediglich in der Berechnung von $[\tilde{v}](\mathbf{q})$ nach Gl. (D.5) den Summanden ($\mathbf{G}$=$\mathbf{0}$) wegzulassen. Die zweite Summe in Gl. (E.4) führen wir genau so aus, wie sie dort notiert ist: Wir bestimmen die entsprechenden Elemente (Kopf und Flügel) der dielektrischen Matrix $\tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$ im Fourierraum (dort ist es ohne Divergenz möglich) und summieren dann zur dielektrischen Matrix im Gauß-Orbital-Raum auf. Es soll nochmals betont werden, daß diese besondere Behandlung nur im Falle $\mathbf{q}\rightarrow\mathbf{0}$ erforderlich ist. Dort erweist sie sich allerdings noch aus einem weiteren Grunde als vorteilhaft. Es stellt sich heraus, daß zur konvergenten Bestimmung des Kopfes $\tilde{\epsilon}_{\mathbf{0},\mathbf{0}}(\mathbf{q})$ der dielektrischen Matrix (im Fourierraum) im Falle $\mathbf{q}\rightarrow\mathbf{0}$ wesentlich mehr reziproke Vektoren $\mathbf{k}$ aus der ersten Brillouin-Zone notwendig sind als für den Rest der Matrix. Diese Anforderung läßt sich nun elegant erfüllen. Wir berechnen diejenigen Beiträge zu den Elementen $[\tilde{\epsilon}]_{\beta,\beta'}(\mathbf{q})$, die der ersten Summe in (E.4) entsprechen, mit relativ wenigen $\mathbf{k}$-Punkten direkt in der Form $[\tilde{v}]S^{-1}[P]S^{-1}[\tilde{v}]$. Für die Berechnung der wenigen Elemente $\tilde{\epsilon}_{\mathbf{G},\mathbf{G}'}(\mathbf{q})$, die in die zweite Summe in (E.4) eingehen, also insbesondere für den Kopf $\tilde{\epsilon}_{\mathbf{0},\mathbf{0}}(\mathbf{q})$ der dielektrischen Matrix, benutzen wir ein dichteres $\mathbf{k}$-Punkt-Schema.

Das zweite Problem, das sich aus der Langreichweitigkeit der Coulomb-Wechselwirkung ergibt, betrifft die Berechnung von Matrixelementen des Selbstenergie-Operators. Da die Realisierung der GW-Näherung ohnehin einen enormen numerischen Aufwand bedeutet, wird man bemüht sein, die Auswertung mit möglichst wenig Aufwand zu einem konvergenten Ergebnis zu führen. Konkret bedeutet das, daß die Auswertung der Brillouinzonen-Integration bei der Bestimmung des Matrixelementes $\langle \mathrm{m}\mathbf{k}|\Sigma|\mathrm{m}'\mathbf{k}\rangle$ gemäß (4.12) mit möglichst wenigen speziellen Punkten durchgeführt werden soll. Um die Divergenz-Problematik in dieser Summe erkennen zu können, stellen wir den Ausdruck hier in Form ebener Wellen dar:

$$\begin{aligned}
\left\langle \psi_{\mathrm{m}\mathbf{k}}|\Sigma(E^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}})|\psi_{\mathrm{m}'\mathbf{k}}\right\rangle &= \frac{4\pi e^2}{V}\sum_{\mathbf{q}}\sum_{\mathbf{G},\mathbf{G}'}\sum_{\mathrm{n}}\frac{1}{|\mathbf{q}+\mathbf{G}|}\frac{1}{|\mathbf{q}+\mathbf{G}|'}F^{\mathrm{mn}}_{\mathbf{G}}(\mathbf{k},\mathbf{q})\left(F^{\mathrm{m}'\mathrm{n}}_{\mathbf{G}'}(\mathbf{k},\mathbf{q})\right)^* \\
&\quad\times\frac{i}{2\pi}\int e^{i\omega 0^+}\frac{\tilde{\epsilon}^{-1}_{\mathbf{G},\mathbf{G}'}(-\mathbf{q},\omega)}{E^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}}-\omega-E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}+i0^+\mathrm{sign}(E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}}-E_F)}d\omega \qquad \text{(E.5)} \\
&\hat{=}\ \sum_{\mathbf{q}}\sum_{\mathbf{G},\mathbf{G}'}\frac{1}{|\mathbf{q}+\mathbf{G}|}\frac{1}{|\mathbf{q}+\mathbf{G}|'}f_{\mathbf{G},\mathbf{G}'}(\mathbf{q}) \qquad \text{(E.6)}
\end{aligned}$$

$$\text{mit}\quad F^{\mathrm{mn}}_{\mathbf{G}}(\mathbf{k},\mathbf{q}) = \int \psi^*_{\mathrm{m}\mathbf{k}}(\mathbf{r})e^{-i(\mathbf{q}+\mathbf{G})\mathbf{r}}\psi_{\mathrm{n},\mathbf{k}+\mathbf{q}}(\mathbf{r})d^3r \quad . \qquad \text{(E.7)}$$

Divergenzen für $\mathbf{q}\to\mathbf{0}$ treten auf, wenn $\mathbf{G}$ oder $\mathbf{G}'$ Null ist. Wir untersuchen hier hauptsächlich den Term $\mathbf{G}=\mathbf{G}'=0$. Die Divergenzen der gemischten Terme (nur $\mathbf{G}=0$ oder nur $\mathbf{G}'=0$) können ähnlich behandelt werden, sind für die Selbstenergie aber unbedeutend. Für $\mathbf{q}\to\mathbf{0}$ verhalten sich die einzelnen Elemente des ($\mathbf{G}=\mathbf{G}'=0$)-Terms in (E.5) wie folgt:

$$F_{\mathbf{0}}^{\mathrm{mn}}(\mathbf{k},\mathbf{q}) \;\to\; \delta_{mn} \tag{E.8}$$

$$E^{\mathrm{LDA}}_{\mathrm{m}'\mathbf{k}} - E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}} \;\to\; 0 \quad \text{für} \quad m' = n \quad . \tag{E.9}$$

Für $m \neq m'$ verschwindet demnach der Term $f_{\mathbf{0},\mathbf{0}'}(\mathbf{0})$. Divergenzen treten nur in den Diagonalelementen $m = m'$ des Selbstenergie-Operators auf. Wir stellen folgendes Ergebnis fest:

$$\frac{1}{q^2} f_{\mathbf{00}}(\mathbf{q}) \;\to\; \frac{1}{q^2}\frac{i}{2\pi}\int e^{i\omega 0^+} \frac{\tilde{\epsilon}^{-1}_{\mathbf{0},\mathbf{0}'}(\mathbf{0},\omega)}{-\omega + i0^+\mathrm{sign}(E^{\mathrm{LDA}}_{\mathrm{n},\mathbf{k}+\mathbf{q}} - E_F)} d\omega \tag{E.10}$$

$$= \frac{1}{q^2} \times \begin{cases} -1 + \dfrac{1}{2}\left(1 - \dfrac{1}{\epsilon_\infty}\right) & \text{für m}\in\text{Val} \\ -\dfrac{1}{2}\left(1 - \dfrac{1}{\epsilon_\infty}\right) & \text{für m}\in\text{Con} \quad . \end{cases} \tag{E.11}$$

Man gelangt daraus zu folgender Vorgehensweise zur Auswertung der Brillouinzonen-Integration:
(i) Man berechne das Integral (E.5) bzw. (4.12) mit einem Abtastraster $\{\mathbf{q}_i\}$ (jeweiliges Gewicht w_i; $\sum w_i=1$), das den Punkt $\mathbf{q}=\mathbf{0}$ nicht enthalten darf.
(ii) Man berücksichtige die Divergenz am Punkt $\mathbf{q}=\mathbf{0}$ durch Addition eines Korrekturterms

$$\left(\frac{V}{(2\pi)^3}\int_{1.\mathrm{BZ}} g(\mathbf{q}) d^3q - \sum_i w_i g(\mathbf{q}_i)\right) \times \begin{cases} -1 + \dfrac{1}{2}\left(1 - \dfrac{1}{\epsilon_\infty}\right) & \text{für m}\in\text{Val} \\ -\dfrac{1}{2}\left(1 - \dfrac{1}{\epsilon_\infty}\right) & \text{für m}\in\text{Con} \quad . \end{cases} \tag{E.12}$$

Die Hilfsfunktion $g(\mathbf{q})$ muß für $\mathbf{q}\to\mathbf{0}$ wie $1/q^2$ divergieren. Die nächstliegende Wahl für die Funktion $g(\mathbf{q})$ lautet $g(\mathbf{q}) = 1/q^2$. Das Integral über die erste Brillouinzone muß dann numerisch ausgewertet werden. Oft erweist es sich als günstiger, für $g(\mathbf{q})$ eine gitterperiodische Funktion zu wählen, die für $\mathbf{q}\to\mathbf{0}$ wie $1/q^2$ divergiert, und deren Integral über die erste Brillouinzone analytisch bekannt ist. Für die kubisch flächenzentrierte Gitterstruktur kann man

$$g(\mathbf{q}) = \frac{(a_0/2)^2}{3 - \cos(\frac{a_0 q_x}{2})\cos(\frac{a_0 q_y}{2}) - \cos(\frac{a_0 q_y}{2})\cos(\frac{a_0 q_z}{2}) - \cos(\frac{a_0 q_x}{2})\cos(\frac{a_0 q_z}{2})} \tag{E.13}$$

wählen [Gyg86] und erhält für das Integral über die erste Brillouin-Zone:

$$\frac{V}{(2\pi)^3}\int\limits_{1.\mathrm{BZ}} g(\mathbf{q})d^3q = \left(\frac{a_0}{2\pi}\right)^2\sqrt{3}\left(K\left(\sin(\frac{\pi}{12})\right)\right)^2 = 4.423758\left(\frac{a_0}{2\pi}\right)^2 \quad . \qquad \text{(E.14)}$$

Dabei ist $K(x)$ das vollständige elliptische Integral erster Gattung.

Bei der Behandlung der (001) Oberflächen tritt eine quaderförmige Superzelle mit den Abmessungen a_x, a_y, a_z auf (O.E. $a_x \leq a_y \leq a_z$). Hier wählen wir die Funktion

$$g(\mathbf{q}) \;=\; \frac{1}{A - w_x cos(a_x q_x) - w_y cos(a_y q_y) - w_y cos(a_z q_z)} \qquad \text{(E.15)}$$

$$\text{mit} \qquad w_j = \frac{2}{q_j^2} \quad \text{und} \quad A = w_x + w_y + w_z \qquad \text{(E.16)}$$

und bestimmen das Integral über die erste Brillouinzone zu

$$\frac{V}{(2\pi)^3}\int\limits_{1.\mathrm{BZ}} g(\mathbf{q})d^3q \;=\; \frac{1}{\pi^2}\frac{1}{\sqrt{w_x w_y}}\int_0^\pi dp_z \eta(p_z) K(\eta(p_z)) \qquad \text{(E.17)}$$

$$\text{mit} \quad \eta(p_z) \;=\; 2\sqrt{\frac{w_x w_y}{(A - w_z\cos p_z)^2 - (w_y - w_x)^2}} \quad . \qquad \text{(E.18)}$$

[$K(x)$: vollständiges elliptisches Integral erster Gattung.] Das verbleibende eindimensionale Integral läßt sich nach einigen Umformungen stabil numerisch berechnen.

Zur analogen Behandlung hexagonaler Brillouinzonen verweisen wir auf die Arbeit von Wenzien et al. [Wen95a].

Zwei Punkte sind bei dieser Divergenzbehandlung noch zu beachten:
(i) In Gl. (E.11) ist nur von einer einzigen makroskopischen dielektrischen Konstanten die Rede. Das ist bei kubischen Systemen korrekt. Bei nichtkubischen Systemen muß der Formalismus erweitert werden [Wen95a], da in der dielektrischen Matrix eine Richtungsabhängigkeit auftritt; der Grenzwertübergang $\mathbf{q}\rightarrow\mathbf{0}$ ergibt für verschiedene Richtungen von $\mathbf{q}$ verschiedene Ergebnisse (z.B. $\epsilon_\parallel \neq \epsilon_\perp$ in hexagonalen Systemen). Die dadurch auftretenden Änderungen in den Selbstenergien sind allerdings im allgemeinen ohne große Bedeutung.
(ii) Wir haben die in Gl. (E.11) auftretende statische dielektrische Konstante mit ϵ_∞ bezeichnet. „Statisch“ bedeutet im Sinne unseres Formalismus: statisch im Vergleich zur Elektronendynamik. Bei ionischen Kristallen treten bei Energien im meV-Bereich Beiträge zur dielektrischen Funktion auf, die von den Phononen herrühren,

und die im allgemeinen eine Verstärkung der dielektrischen Abschirmung bewirken: $\epsilon_0 > \epsilon_\infty$. Diese Beiträge sind für den Selbstenergie-Operator nicht relevant; sie treten im Rahmen der Born-Oppenheimer Näherung auch gar nicht auf. In unserem Formalismus haben wir die Konstante ϵ_∞ zu verwenden. Die Bezeichnung dieser Konstante als „Konstante bei unendlich hoher Frequenz“ (∞) ist historisch zu verstehen; sie bezieht sich auf Experimente und Überlegungen, bei denen die Phononen im Vordergrund stehen.

Bevor wir das gesamte Schema auf Oberflächenzustände anwenden können, bedarf es noch einiger Erweiterungen. Der Selbstenergie-Operator ist ein kurzreichweitiger Operator. Bei der Berechnung seines Erwartungswertes für einen lokalisierten Oberflächenzustand ist die makroskopische dielektrische Konstante, die einem langreichweitigen Mittelungsprozeß über das gesamte System entspricht, nicht relevant. Wichtiger ist vielmehr eine ortsabhängige dielektrische Konstante $\epsilon_\infty(\mathbf{r})$ als Ergebnis einer kurzreichweitigen Mittelung. Im Volumensystem stellt das keinen Unterschied dar, da bereits die Mittelung über eine Einheitszelle auf die gewöhnliche makroskopische Konstante führt. An der Oberfläche wird man jedoch einen Übergang der ortsabhängigen Konstante vom Volumenwert $\epsilon_\infty^{\text{bulk}}$ auf den Vakuumwert $\epsilon_\infty^{\text{Vak}}$=1 beobachten. Dieser Übergang ist für einen lokalisierten Zustand von Bedeutung.

In ihrer ursprünglichen Definition (3.19) besitzt die dielektrische Konstante eines elektronischen Systems prinzipiell keine Ortsabhängigkeit. Sie beschreibt die über den ganzen Raum gemittelte Reaktion des Systems auf eine äußere Störung mit (im Grenzwert) unendlicher Wellenlänge. Ein sinnvoller kurzreichweitiger ortsabhängiger Mittelungsprozeß kann in der Form

$$\epsilon_\infty^{-1}(\mathbf{r}) = \int\limits_{B(\mathbf{r})} d^3r' \int d^3r'' \epsilon^{-1}(\mathbf{r}', \mathbf{r}'') \tag{E.19}$$

vorgenommen werden. Dabei ist $B(\mathbf{r})$ ein geeigneter Bereich um $\mathbf{r}$. Diese (inverse) dielektrische Konstante beschreibt die lokal gemittelte Reaktion des System auf eine Störung großer Wellenlänge. Der Bereich B sollte größer als die mikroskopische Oszillationslänge der elektronischen Eigenschaften, also größer als einige Bindungslängen sein. Bezogen auf die von uns betrachteten Oberflächen bedeutet das, daß parallel zur Oberfläche die gewöhnliche Mittelung bezüglich $\mathbf{r}_\parallel$ und $\mathbf{r}'_\parallel$ vorgenommen werden kann:

$$\epsilon_\infty^{-1}(z', z'') \;=\; \int dx'dy'dx''dy'' \epsilon^{-1}(\mathbf{r}', \mathbf{r}'') \tag{E.20}$$

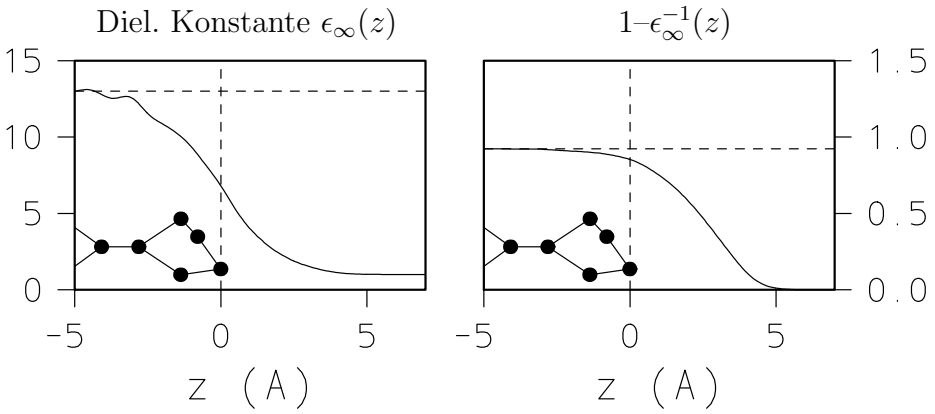


Abbildung E.1: *Zur z-Abhängigkeit der makroskopischen Abschirmung an einer Halbleiteroberfläche (konkret: Si(001)-(2×1)). Links: die dielektrische Konstante (13.0 im Bulk) fällt zum Vakuum hin auf 1 ab. An der Oberfläche (definiert durch das oberste Si Atom) beträgt sie 6.8. Rechts: die für die Selbstenergie relevante Größe $1–\epsilon_\infty^{-1}(z)$ (0.92 im Bulk) fällt zum Vakuum hin auf 0 ab. An der Oberfläche beträgt sie 0.85. Weitere Erläuterungen siehe Text.*

$$\text{bzw.} \quad \epsilon_\infty^{-1}(G'_z, G''_z) = \epsilon^{-1}_{\mathbf{G}',\mathbf{G}''}(\mathbf{q})\delta_{G'_x,0}\delta_{G'_y,0}\delta_{G''_x,0}\delta_{G''_y,0} \quad . \tag{E.21}$$

Anschließend erfolgt die Mittelung über z'' und die Bereichsintegration bzgl. z':

$$\epsilon_\infty^{-1}(z) = \int_{z-c}^{z+c} dz' \int dz'' \epsilon_\infty^{-1}(z', z'') \tag{E.22}$$

$$= \sum_{G_z} e^{iG_z z} \frac{sin(G_z c)}{G_z c} \epsilon_\infty^{-1}(G_z, G'_z)\delta_{G'_z,0} \quad , \tag{E.23}$$

$$\epsilon_\infty(z) = \frac{1}{\epsilon_\infty^{-1}(z)} \quad . \tag{E.24}$$

Falls man für c die Länge der Einheitszelle wählt, so geht dieser Mittelungsprozeß in die gewöhnliche makroskopische dielektrische Konstante über. Kleinere Werte für c führen zu dem räumlichen Übergang der Konstante vom Volumenwert zum Vakuumwert. Sinnvolle Werte für c sind einige Å. Das Ergebnis hängt nur sehr wenig vom konkreten Wert für c ab. In Abb. E.1 stellen wir das Ergebnis solch einer Mittelung für die Si(001)-(2×1) Oberfläche vor. Sie beruht auf der in RPA bestimmten statischen dielektrischen Matrix $[\epsilon](\mathbf{q})$. [Aufgrund der endlichen Slabdicke fällt die daraus berechnete dielektrische Konstante in der Slabmitte geringer aus als im Bulksystem [Wan94]. Wir haben daher das in der Slabmitte bestimmte $\epsilon_\infty(z)$ auf den

Volumenwert ϵ_∞=13.0 skaliert, den wir in einer Volumenrechnung erhalten.] Obwohl die dielektrische Konstante in der Oberflächenebene bereits auf die Hälfte des Volumenwertes gesunken ist, fällt die für die Selbstenergie relevante Größe $1–\epsilon_\infty^{-1}(z)$ nur sehr schwach ab und beträgt an der Oberfläche noch 93% des Volumenwertes. Die Dangling-Bonds sind zum überwiegenden Teil in diesem Oberflächenbereich lokalisiert. Die Divergenz-Korrektur (E.11) zu ihrer Selbstenergie fällt daher etwas anders aus als für Volumenzustände; sie ist für besetzte Zustände etwas niedriger und für unbesetzte Zustände etwas höher. Wir berücksichtigen das, indem wir die effektive dielektrische Konstante für diese Zustände entsprechend modifizieren. Die Bestimmung der effektiven dielektrischen Konstanten für einen Oberflächenzustand weist eine leichte Unsicherheit auf, bedingt durch die etwas willkürliche Mittelungsvorschrift und durch die Bestimmung der dielektrischen Matrix im Slabsystem. Die entsprechende verbleibende Unsicherheit in den Selbstenergien der Oberflächen-Zustände schätzen wir auf etwa 0.05 eV.

Anhang F

Parameterbestimmung in Funktionen mit Polstruktur

In dieser Arbeit treten Funktionen auf, deren Energie-Abhängigkeit eine polartige Struktur aufweist. Das in Kap. 4 vorgestellte Multi Plasmon-Pol Modell dient zur Darstellung der dynamischen dielektrischen Matrix. In Kap. 14 hatten wir ein einfaches Modell für die Selbstenergie eines elektronischen Zustands verwendet, das ebenfalls eine polartige Struktur besitzt. In beiden Fällen war das jeweilige Modell mit Parametern konzipiert worden, die so zu wählen sind, daß das Modell für bestimmte Testfrequenzen die explizit berechnete Funktion reproduziert.

Wir betrachten das Problem zunächst in allgemeiner Form. Gegeben sei eine komplexwertige Funktion $f(\omega)$, die durch ein Modell der Form

$$f(\omega) \quad \doteq \quad a + \sum_{j=1}^{n} \frac{z^{(j)}}{\omega - \omega^{(j)}} \quad . \tag{F.1}$$

approximiert werden soll. Die Frequenz ω kann ebenfalls komplex sein. Um die $(2n{+}1)$ Parameter in diesem Modell zu berechnen, ist die Kenntnis der Funktionswerte für $(2n{+}1)$ Testfrequenzen Ω_k erforderlich. Daraus ergibt sich ein $(2n{+}1)$-dimensionales, nichtlineares Gleichungssystem:

$$\begin{aligned} a + \sum_{j=1}^{n} \frac{z^{(j)}}{\Omega_1 - \omega^{(j)}} \quad &= \quad f(\Omega_1) \\ &\ldots \\ a + \sum_{j=1}^{n} \frac{z^{(j)}}{\Omega_{2n+1} - \omega^{(j)}} \quad &= \quad f(\Omega_{2n+1}) \quad . \end{aligned}$$

Mit der Definition neuer Funktionen

$$F_k(\mathbf{x}) := a + \sum_{j=1}^{n} \frac{z^{(j)}}{\Omega_k - \omega^{(j)}} \quad , \tag{F.2}$$

wobei $\mathbf{x} = (a, z^{(1)}, ..., z^{(n)}, \omega^{(1)}, ..., \omega^{(n)})$ die Parameter des Modells für f umfaßt, lautet dieses Gleichungssystem folgendermaßen:

$$F_k(\mathbf{x}) = f(\Omega_k) \quad , \quad k = 1, ..., 2n+1 \quad . \tag{F.3}$$

Wir suchen nun eine Lösung $\mathbf{x}$ für dieses Gleichungssystem. Dazu betrachten wir das Newton'sche Iterationsverfahren. Ausgehend von einer Startkonfiguration $\mathbf{x}^{(0)}$ bestimmt man iterativ neue Konfigurationen $\mathbf{x}^{(l)}$ für den Parametersatz nach der Vorschrift

$$\mathbf{x}^{(l+1)} := \mathbf{x}^{(l)} - \left(\frac{\partial F_k(\mathbf{x}^{(l)})}{\partial x_i} \right)^{-1} \mathbf{F}(\mathbf{x}^{(l)}) \quad . \tag{F.4}$$

Das Verfahren ist auf unser Problem zunächst nicht anwendbar, denn die Iterationsvorschrift ist nicht kontrahierend. Wegen der Polstruktur der Ausgangsfunktionen gilt für die in (F.4) auftretenden Ableitungen nach den Polfrequenzen:

$$\frac{\partial F_k}{\partial \omega^{(j)}} = \frac{z^{(j)}}{(\Omega_k - \omega^{(j)})^2} \quad . \tag{F.5}$$

Wenn die Frequenzen $\omega^{(j)}$ sehr groß werden, ist die gesamte Ableitungsmatrix nicht mehr gut konditioniert, und ihre Inverse divergiert quadratisch in $|\omega^{(j)}|$, also wie $|\omega^{(j)}|^2$. Da gleichzeitig die Funktionswerte F_k nur wie $1/|\omega^{(j)}|$ gegen Null konvergieren, werden die Parameter $\omega^{(j)}$ durch die Iterationsvorschrift (F.4) beliebig groß. Diese Schwierigkeit umgehen wir, indem wir Gl. (F.3) beidseitig mit einem Faktor $p(\mathbf{x})$ multiplizieren, der sicherstellt, daß der gesamte Ausdruck für $\omega^{(j)} \to \infty$ gegen ∞ divergiert. Dadurch wird die Divergenz der Parameter $\mathbf{x}$ in der Iterationsvorschrift (F.4) verhindert. Der Faktor $p(\mathbf{x})$ darf allerdings nirgends Null werden, da dadurch das Gleichungssystem (F.3) zusätzliche, unphysikalische Lösungen aufweisen würde. Wir wählen für diesen Faktor das Produkt

$$p(\mathbf{x}) = \prod_{j=1}^{n} (|\omega^{(j)}|^2 + 1) > 0 \tag{F.6}$$

Ein unumgänglicher Nachteil einer solchen Wahl besteht darin, daß dieser Term $p(\mathbf{x})$ im Komplexen nicht analytisch ist und daher nicht nach $\omega^{(j)}$ differenziert werden kann. Um das Newton-Verfahren anwenden zu können, muß das gesamte Problem

der Parameterbestimmung in Real- und Imaginärteile zerlegt werden. Dadurch verdoppelt sich sein Dimension auf $(4n+2)$. Das zu lösende *reelle* Gleichungssystem lautet nunmehr

$$\begin{array}{rcl} \Re G_k(\mathbf{y}) & = & \Re g(\Omega_k) \\ \Im G_k(\mathbf{y}) & = & \Im g(\Omega_k) \end{array} \quad , \quad k = 1, ..., 2n+1 \quad . \tag{F.7}$$

mit den Größen

$$\mathbf{y} = (\Re \mathbf{x}, \Im \mathbf{x}) \tag{F.8}$$

$$G_k(\mathbf{y}) = p(\mathbf{x}) F_k(\mathbf{y}) \tag{F.9}$$

$$g(\Omega_k) = p(\mathbf{x}) f(\Omega_k) \quad . \tag{F.10}$$

Wir lösen es mit dem Newton'schen Iterationsverfahren. Die Lösung für $\mathbf{y}$, die wir erhalten, enthält die Real- und Imaginärteile der gesuchten Parameter.

Die Energieabhängigkeit der Selbstenergie, die wir in Kap. 14 betrachtet haben, besitzt genau die Struktur, für die wir hier eine Lösungsmöglichkeit vorgestellt haben. Die Behandlung der Multi Plasmon-Pol Modelle erfolgt in analoger Form. Wir haben ihre Frequenzabhängigkeit mit Hilfe von Frequenzquadraten formuliert (siehe Gl. 4.15). Man behandle daher zunächst diese Quadrate als Parametergrößen und ziehe anschließend aus den berechneten Parametern die Wurzel. Im Unterschied zur Selbstenergie tritt bei den PPM ein Parameter weniger auf, da die Konstante a beim PPM als 1 bzw. als $S_{\beta\beta'}$ festgelegt ist (vgl. Gl. (4.15)). Daher ist auch eine Testfrequenz weniger zur Parameterbestimmung notwendig. Wenn bei der Festlegung der Parameter des PPM die Johnson'sche f-Summenregel berücksichtigt werden soll, so reduziert sich die Zahl der Testfrequenzen nochmals um Eins. Im Gleichungssystem (F.3) wird dabei eine der Gleichungen durch die f-Summenregel ersetzt.

Literaturverzeichnis

[And95] G. W. Anderson, M. C. Hanf, P. R. Norton, Z. H. Lu und M. J. Graham, Appl. Phys. Lett. **66**, 1123 (1995).

[App78] J. A. Appelbaum, G. A. Baraff, D. R. Hamann, H. D. Hagstrum und T. Sakurai, Surf. Sci. **70**, 654 (1978).

[Ary92] F. Aryasetiawan, Phys. Rev. B **46**, 13 051 (1992).

[Ary95] F. Aryasetiawan und O. Gunnarsson, Phys. Rev. Lett. **74**, 3221 (1995).

[Ary96] F. Aryasetiawan und O. Gunnarsson, unveröffentlicht.

[Bac82] G. B. Bachelet, D. R. Hamann und M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[Bac95] W. H. Backes, P. A. Bobbert und W. van Haeringen, Phys. Rev. B **51**, 4950 (1995).

[Bec81] F. Bechstedt, R. Enderlein und R. Wischnewski, phys. stat. sol. (b) **107**, 637 (1981).

[Bec94] F. Bechstedt, M. Fiedler, C. Kreß und R. Del Sole, Phys. Rev. B **49**, 7357 (1994).

[Bla94a] X. Blase, X. Zhu und S. G. Louie, Phys. Rev. B **49**, 4973 (1994).

[Bla94b] X. Blase, A. J. R. da Silva, X. Zhu und S. G. Louie, Phys. Rev. B **50**, 8102 (1994).

[Bla95] X. Blase, A. Rubio, S. G. Louie und M. L. Cohen, Phys. Rev. B **51**, 6868 (1995).

[Blo72] C. Blomberg und B. Bergersen, Can. J. Phys. **50**, 2286 (1972).

[Bob94] P. A. Bobbert und W. van Haeringen, Phys. Rev. B **49**, 10 326 (1994).

[Bru93] T. Brudevoll, D. S. Citrin, M. Cardona und N. E. Christensen, Phys. Rev. B **48**, 8629 (1993).

[Bul95] E. L. Bullock, R. Gunnella, L. Patthey, T. Abukawa, S. Kono, C. R. Natoli und L. S. O. Johansson, Phys. Rev. Lett. **74**, 2756 (1995).

[Bul96] E. L. Bullock, R. Gunnella, C. R. Natoli, R. I. G. Uhrberg und L. S. O. Johansson, unveröffentlicht

[Büs94] H. Büscher, Dissertation (Münster, 1994).

[Car65] M. Cardona, M. Weinstein und G. A. Wolff, Phys. Rev. **140**, A 633 (1965).

[Cep80] D. M. Ceperley und B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[Cha73] D. J. Chadi und M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[Cha79] D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).

[Cha83] Y. J. Chabal, S. B. Christman, E. E. Chaban und M. T. Yin, J. Vac. Sci. Technol. A **1**, 1241 (1983).

[Cha93] J. P. A. Charlesworth, R. W. Godby und R. J. Needs, Phys. Rev. Lett. **70**, 1685 (1993)

[Che76] J. R. Chelikowsky und M. L. Cohen, Phys. Rev. B **14**, 556 (1976)

[Che80] C. H. Chen, A. E. Meixner und B. M. Kincaid, Phys. Rev. Lett. **44**, 951 (1980).

[Cho93] K. Cho und J. D. Joannopoulos, Phys. Rev. Lett. **71**, 1387 (1993).

[Cul86] R. J. Culbertson, Y. Kuk und L. C. Feldman, Surf. Sci. **167**, 127 (1986).

[Dab92] J. Dabrowski und M. Scheffler, Appl. Surf. Sci. **56**, 15 (1992).

[Din96] S. A. Ding, G. Neuhold, J. H. Weaver, P. Häberle, K. Horn, O. Brandt, H. Yang und K. Ploog, J. Vac. Sci. Technol. A**14**, 819 (1996).

[Dum96] P. Dumas, private Mitteilung.

[Eng90] G. E. Engel und R. J. Needs, Phys. Rev. B **41**, 7876 (1990).

[Fal91] C. Falter, M. Klenner und W. Ludwig, phys. stat. sol. (b) **167**, 85 (1991).

[Fal92] C. Falter, private Mitteilung.

[Fio93] V. Fiorentini, M. Methfessel und M. Scheffler, Phys. Rev. B **47**, 13 353 (1993).

[Fle96] A. Fleszar, private Mitteilung.

[Fon93] E. Fontes, J. R. Patel und F. Comin, Phys. Rev. Lett. **70**, 2790 (1993).

[Ful91] P. Fulde, *Electronic Correlation in Molecules and Solids* (Springer, Berlin 1991).

[God88] R. W. Godby, M. Schlüter und L. J. Sham, Phys. Rev. B **37**, 10 159 (1988).

[Gro95] H. J. de Groot, P. A. Bobbert und W. van Haeringen, Phys. Rev. B **52**, 11 000 (1995).

[Gry95] J. Gryko und O. F. Sankey, Phys. Rev. B **51**, 7295 (1995).

[Gyg86] F. Gygi und A. Baldereschi, Phys. Rev. B **34**, 4405 (1986).

[Gyg93] F. Gygi, Phys. Rev. B **48**, 11 692 (1993).

[Ham79] D. R. Hamann, M. Schlüter und C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[Ham86] R. J. Hamers, R. M. Tromp und J. E. Demuth, Phys. Rev. B **34**, 5343 (1986).

[Ham89] R. J. Hamers und U. K. Köhler, J. Vac. Sci. Technol. A **7**, 2854 (1989).

[Ham91] D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).

[Han87] D. Haneman, Rep. Prog. Phys. **50**, 1045 (1987).

[Hed65] L. Hedin, Phys. Rev. **139**, A796 (1965).

[Hed69] L. Hedin und S. Lundqvist, in *Solid State Physics, Advances in Research and Application*, herausgegeben von F. Seitz, D. Turnbull und H. Ehrenreich (Academic, New York, 1969), Band 23, S. 1.

[Hen94] M. Henzler und W. Göpel, *Oberflächenphysik des Festkörpers*, 2. Auflage (Teubner, Stuttgart 1994).

[Hil95] G. Hilgers, M. Potthoff, N. Müller, U. Heinzmann, L. Haunert, J. Braun und G. Borstel, Phys. Rev. B **52**, 14 859 (1995).

[Him80] F. J. Himpsel, P. Heimann, T.-C. Chiang und D. E. Eastman, Phys. Rev. Lett. **45**, 1112 (1980).

[Him79] F. J. Himpsel und D. E. Eastman, J. Vac. Sci. Technol. **16**, 1297 (1979).

[Höc83] H. Höchst und J. Hernández-Calderón, Surf. Sci. **126**, 25 (1983).

[Hof93] P. Hofmann, K. Horn, A. M. Bradshaw, R. L. Johnson, D. Fuchs und M. Cardona, Phys. Rev. B **47**, 1639 (1993).

[Hoh64] P. Hohenberg und W. Kohn, Phys. Rev. **136**, B864 (1964).

[Hol84] B.W. Holland, C. B. Duke und A. Paton, Surf. Sci. **140**, L269 (1984).

[Hor87] P. Horsch, W. von der Linden und W.-D. Lukas, Solid State Comm. **62**, 359 (1987).

[Hor96] K. Horn, private Mitteilung.

[Hot91] R. Hott, Phys. Rev. B **44**, 1057 (1991).

[Hri93] K. Hricovini, R. Günther, P. Thiry, A. Taleb-Ibrahimi, G. Indlekofer, J. E. Bonnet, P. Dumas, Y. Petroff, X. Blase, X. Zhu, S. G. Louie, Y. J. Chabal und P. A. Thiry, Phys. Rev. Lett. **70**, 1992 (1993).

[Hua93] M.-Z. Huang und W. Y. Ching, Phys. Rev. B **47**, 9449 (1993).

[Hyb86] M. S. Hybertsen und S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[Hyb87a] M. S. Hybertsen und S. G. Louie, Phys. Rev. Lett. **58**, 1551 (1987).

[Hyb87b] M. S. Hybertsen und M. Schlüter, Phys. Rev. B **36**, 9683 (1987).

[Hyb88a] M. S. Hybertsen und S. G. Louie, Phys. Rev. B **37**, 2733 (1988).

[Hyb88b] M. S. Hybertsen und S. G. Louie, Phys. Rev. B **38**, 4033 (1988).

[Iba95] H. Ibach und H. Lüth, *Festkörperphysik*, 4. Auflage (Springer, Berlin 1995).

[Ihm83] J. Ihm, D. H. Lee, J. D. Joannopoulos und J. J. Xiong, Phys. Rev. Lett. **51**, 1872 (1983).

[Joh74] D. L. Johnson, Phys. Rev. B **9**, 4475 (1974).

[Joh90] L. S. O. Johansson, R. I. G. Uhrberg, P. Mårtensson und G. V. Hansson, Phys. Rev. B **42**, 1305 (1990).

[Joh92] L. S. O. Johansson und B. Reihl, Surf. Sci. **269/270**, 810 (1992).

[Kev84] S. D. Kevan und N. G. Stoffel, Phys. Rev. Lett. **53**, 702 (1984).

[Kev85] S. D. Kevan, Phys. Rev. B **32**, 2344 (1985).

[Kip92] L. Kipp, R. Manzke und M. Skibowski, Surf. Sci. **269/270**, 854 (1992).

[Kip95] L. Kipp, R. Manzke und M. Skibowski, Solid State Comm. **93**, 603 (1995).

[Kip96] L. Kipp, private Mitteilung.

[Kle82] L. Kleinman und D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[Koh65] W. Kohn und L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[Kre93] C. Kreß, M. Fiedler und F. Bechstedt, *Proceedings of the Fourth International Conference on the Formation of Semiconductor Interfaces* (Jülich, 1993), herausgegeben von B. Lengeler, H. Lüth, W. Mönch und J. Pollmann (World Scientific, Singapur, 1994), S. 19.

[Kre94] C. Kreß, M. Fiedler, W. G. Schmidt und F. Bechstedt, Phys. Rev. B **50**, 17 697 (1994).

[Kre95] C. Kreß, M. Fiedler, W. G. Schmidt und F. Bechstedt, Surf. Sci. **333**, B1152 (1995).

[Krü86] P. Krüger, A. Mazur, J. Pollmann und G. Wolfgarten, Phys. Rev. Lett. **57**, 1468 (1986).

[Krü90] P. Krüger und J. Pollmann, Phys. Rev. Lett. **64**, 1808 (1990).

[Krü91] P. Krüger und J. Pollmann, Prog. Surf. Sci. **35**, 3 (1991).

[Krü93] P. Krüger und J. Pollmann, Phys. Rev. B **47**, 1898 (1993).

[Krü94] P. Krüger und J. Pollmann, Appl. Phys. A**59**, 487 (1994).

[Krü95] P. Krüger und J. Pollmann, Phys. Rev. Lett. **74**, 1155 (1995).

[KrüPr] P. Krüger, private Mitteilungen.

[Kub87] J. A. Kubby, J. E. Griffith, R. S. Becker und J. S. Vickers, Phys. Rev. B **36**, 6079 (1987).

[Lam87] W. R. Lambert, P. L. Trevor, M. J. Cardillo, A. Sakai und D. R. Hamann, Phys. Rev. B **35**, 8055 (1987).

[Lan50] *Zahlenwerte und Funktionen aus Naturwissenschaft und Technik*, Landolt-Börnstein, Neue Ausgabe, Gruppe I, Band 1, herausgegeben von A. Eucken (Springer, Berlin, 1950).

[Lan82] *Zahlenwerte und Funktionen aus Naturwissenschaft und Technik*, Landolt-Börnstein, Neue Ausgabe, Gruppe III, Band 17a, 17b und 22a, herausgegeben von K.-H. Hellwege und O. Madelung (Springer, Berlin, 1982).

[Lan89] *Zahlenwerte und Funktionen aus Naturwissenschaft und Technik*, Landolt-

Börnstein, Neue Ausgabe, Gruppe III, Band 23a, herausgegeben von A. Goldmann und E.-E. Koch (Springer, Berlin, 1989).

[Lan90] E. Landemark, R. I. G. Uhrberg, P. Krüger und J. Pollmann, Surf. Sci. Lett. **236**, L359 (1990).

[Lan92] E. Landemark, C. J. Karlsson, Y.-C. Chao und R. I. G. Uhrberg, Phys. Rev. Lett. **69**, 1588 (1992).

[Lan94] E. Landemark, C. J. Karlsson, L. S. O. Johansson und R. I. G. Uhrberg, Phys. Rev. B **49**, 16 523 (1994).

[Lei92] T. Lei, T. D. Moustakas, R. J. Graham, Y. He und S. Berkowitz, J. Appl. Phys. **71**, 4933 (1992).

[Leu88] K. T. Leung, L. J. Terminello, J. Hussain, X. S. Zhang, T. Hayashi und D. A. Shirley, Phys. Rev. B **38**, 8241 (1988).

[Lev82] Z. H. Levine und S. G. Louie, Phys. Rev. B **25**, 6310 (1982).

[Ley74] L. Ley, R. A. Pollak, F. R. McFeely, S. P. Kowalczyk und D. A. Shirley, Phys. Rev. B **9**, 600 (1974).

[Lin88] W. von der Linden und P. Horsch, Phys. Rev. B **37**, 8351 (1988).

[Liz95] A. Lizón-Nordström und F. Yndurain, Solid State Comm. **94**, 335 (1995).

[Lou82] S. G. Louie, S. Froyen und M. L. Cohen. Phys. Rev. B **26**, 1738 (1982).

[Luc93] C. A. Lucas, C. S. Dower, D. F. McMorrow, G. C. L. Wong, F. J. Lamelas und P. H. Fuoss, Phys. Rev. B **47**, 10 375 (1993).

[Mah89] G. D. Mahan und B. E. Sernelius, Phys. Rev. Lett. **62**, 2718 (1989).

[Mar91] J. L. Martins, N. Troullier und S.-H. Wei, Phys. Rev. B **43**, 2213 (1991).

[Mön81] W. Mönch, P. Koke und S. Krueger, J. Vac. Sci. Technol. **19**, 313 (1981).

[Mön93] W. Mönch, *Semiconductor Surfaces and Interfaces*, Springer Series in Surface Sciences 26 (Springer, Berlin 1993).

[Mun95] A. W. Munz, C. Ziegler und W. Göpel, Phys. Rev. Lett. **74**, 2244 (1995).

[Nee88] M. Needels, M. C. Payne und J. D. Joannopoulos, Phys. Rev. B **38**, 5543 (1988).

[New93] K. Newstead, A. W. Robinson, S. d'Addato, A. Patchett, N. P. Prince, R. McGrath, R. Whittle, E. Dudzik und I. T. McGovern, Surf. Sci. **287/288**, 317 (1993).

[Nor87] J. E. Northrup, M. S. Hybertsen und S. G. Louie, Phys. Rev. Lett. **59**, 819 (1987).

[Nor89] J. E. Northrup, M. S. Hybertsen und S. G. Louie, Phys. Rev. B **39**, 8198 (1989).

[Nor91] J. E. Northrup, M. S. Hybertsen und S. G. Louie, Phys. Rev. Lett. **66**, 500 (1991).

[Nor93] J. E. Northrup, Phys. Rev. B **47**, 10 032 (1993).

[Oba86] S. Obara und A. Saika, J. Chem. Phys. **84**, 3963 (1986).

[Oni95] G. Onida, L. Reining, R. W. Godby, R. Del Sole und W. Andreoni, Phys. Rev. Lett. **75**, 818 (1995).

[Ort93] J. E. Ortega und F. J. Himpsel, Phys. Rev. B **47**, 2130 (1993).

[Pal94] M. Palumno, L. Reining, R. W. Godby, C. M. Bertoni und N. Börnsen, Europhys. Lett. **26**, 607 (1994).

[Pap86] L. Papagno, X. Y. Shen, J. Anderson, G. Schirripa Spagnolo und G. J. Lapeyre, Phys. Rev. B**34**, 7188 (1986).

[Pan95] O. Pankratov und M. Scheffler, Phys. Rev. Lett. **75**, 701 (1995).

[Pas96] A. Pasquarello, M. S. Hybertsen und R. Car, Phys. Rev. B **53**, 10 942 (1996).

[Peh93] E. Pehlke und M. Scheffler, Phys. Rev. Lett. **71**, 2338 (1993).

[Per81] J. P. Perdew und A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[Pic84] W. E. Pickett und C. S. Wang, Phys. Rev. B **30**, 4719 (1984).

[Pol80] J. Pollmann, in *Festkörperprobleme, Advances in Solid State Physics* (Vieweg, Braunschweig 1980), Band XX, S. 117.

[Pol87] J. Pollmann, P. Krüger und A. Mazur, J. Vac. Sci. Technol. B**5**, 945 (1987).

[Pol92] J. Pollmann, P. Krüger und A. Mazur, Appl. Surf. Sci. **56-58**, 193 (1992).

[Pol95] T. Pollehn, private Mitteilung.

[Pop78] T. D. Poppendieck, T. C. Ngoc und M. B. Webb, Surf. Sci. **75**, 287 (1978).

[Pow91] J. M. Powers, A. Wander, P. J. Rous, M. A. van Hove und G. A. Somorjai, Phys. Rev. B **44**, 11 159 (1991).

[Ram95] A. Ramstad, G. Brocks und P. J. Kelly, Phys. Rev. B **51**, 14 504 (1995).

[Rie95] M. M. Rieger und P. Vogl, Phys. Rev. B**52**, 16 567 (1995).

[Roh93a] M. Rohlfing, Diplomarbeit (1993).

[Roh93b] M. Rohlfing, P. Krüger und J. Pollmann, Phys. Rev. B **48**, 17 791 (1993).

[Roj95] H. N. Rojas, R. W. Godby und R. J. Needs, Phys. Rev. Lett. **74**, 1827 (1995).

[Ros92] R. Rossmann, H. L. Meyerheim, V. Jahns, J. Wever, W. Moritz, D. Wolf, D. Dornisch und H. Schulz, Surf. Sci. **279**, 199 (1992).

[Ros93] U. Rossow, T. Werninghaus, D. R. T. Zahn, W. Richter und K. Horn, Thin Solid Films **233**, 176 (1993).

[Rub93] A. Rubio, J. L. Corkill, M. L. Cohen, E. L. Shirley und S. G. Louie, Phys. Rev. B **48**, 11 810 (1993).

[Sab93] M. Sabisch, Diplomarbeit (Münster, 1993).

[Sab94] M. Sabisch, private Mitteilung.

[Sab95] M. Sabisch, P. Krüger und J. Pollmann, Phys. Rev. B**51**, 13 367 (1995).

[Sab96a] M. Sabisch, P. Krüger, A. Mazur, M. Rohlfing und J. Pollmann, Phys. Rev. B**53**, 13 121 (1996).

[Sab96b] M. Sabisch, Dissertation (Münster, 1996).

[Sch75] M. Schlüter, J. R. Chelikowsky, S. G. Louie und M. L. Cohen, Phys. Rev. B **12**, 4200 (1975).

[Sch93a] P. Schröer, P. Krüger und J. Pollmann, Phys. Rev. B **47**, 6971 (1993).

[Sch93b] P. Schröer, P. Krüger und J. Pollmann, Phys. Rev. B **48**, 18 264 (1993).

[Shi92] E. L. Shirley, X. Zhu und S. G. Louie, Phys. Rev. Lett. **69**, 2955 (1992).

[Shi93a] E. L. Shirley und S. G. Louie, Phys. Rev. Lett. **71**, 133 (1993).

[Shi93b] E. L. Shirley und R. M. Martin, Phys. Rev. B**47**, 15 404 und 15 413 (1993).

[Shk95] A. I. Shkrebtii, R. Di Felice, C. M. Bertoni und R. Del Sole, Phys. Rev. B **51**, 11 201 (1995).

[Ski94] M. Skibowski und L. Kipp, J. Electron Spectrosc. Relat. Phenom. **68**, 77 (1994).

[Sol94] R. Del Sole, L. Reining und R. W. Godby, Phys. Rev. B **49**, 8024 (1994).

[Spi94] L. Spiess, A. J. Freeman und P. Soukiassian, Phys. Rev. B **50**, 2249 (1994).

[Sta96] A. P. J. Stampfl, P. Hofmann, O. Schaff und A. M. Bradshaw, unveröffentlicht.

[Sto83] N. G. Stoffel, Phys. Rev. B **28**, 3306 (1983).

[Str82] G. Strinati, Phys. Rev. Lett. **49**, 1519 (1982).

[Stu90] R. Stumpf, X. Gonze und M. Scheffler, Fritz-Haber-Institut Research Report, Berlin 1990.

[Sur95] M. P. Surh, H. Chacham und S. G. Louie, Phys. Rev. B **51**, 7464 (1995).

[Sva90] A. Svane und O. Gunnarson, Phys. Rev. Lett. **65**, 1148 (1990).

[Tab87] T. Tabata, T. Aruga und Y. Murata, Surf. Sci. **179**, L63 (1987).

[Ter85] J. Tersoff und D. R. Hamann, Phys. Rev. B **31**, 805 (1985).

[Tro83] R. M. Tromp, R. G. Smeenk, F. W. Saris und D. J. Chadi, Surf. Sci. **133**, 137 (1983).

[Uhr81] R. I. G. Uhrberg, G. V. Hansson, J. M. Nicholls und S. A. Flodström, Phys. Rev. B **24**, 4684 (1981).

[Uhr87] R. I. G. Uhrberg, R. D. Bringans, M. A. Olmstead, R. Z. Bachrach und J. E. Northrup, Phys. Rev. B **35**, 3945 (1987).

[Ver95] C. Verdozzi, R. W. Godby und S. Holloway, Phys. Rev. Lett. **74**, 2327 (1995).

[Vog95] D. Vogel, P. Krüger und J. Pollmann, Phys. Rev. B **52**, R14 316 (1995).

[Vog96] D. Vogel, P. Krüger und J. Pollmann, Phys. Rev. B **54**, 5495 (1996).

[Wan81] C. S. Wang und B. M. Klein, Phys. Rev. B **24**, 3393 (1981).

[Wan94] L.-W. Wang und A. Zunger, Phys. Rev. Lett. **73**, 1039 (1994).

[Wen95a] B. Wenzien, G. Cappellini und F. Bechstedt, Phys. Rev. B **51**, 14 701 (1995).

[Wen95b] B. Wenzien, P. Käckell, F. Bechstedt und G. Cappellini, Phys. Rev. B **52**, 10 897 (1995).

[Wes87] T. Weser, A. Bogen, B. Konrad, R. D. Schnell, C. A. Schug und W. Steinmann, Phys. Rev. B **35**, 8184 (1987).

[Wes88] T. Weser, A. Bogen, B. Konrad, R. D. Schnell, C. A. Schug, W. Moritz und W. Steinmann, Surf. Sci. **201**, 245 (1988).

[Wig37] E. Wigner, Phys. Rev. **46**, 1002 (1937).

[Woo94] D. P. Woodruff und T. A. Delchar, *Modern Techniques of Surface Science*, 2. Auflage (Cambridge University Press, Cambridge 1994).

[Wri94] A. F. Wright und J. S. Nelson, Phys. Rev. B **50**, 2159 (1994).

[Wür95] K. Würde, private Mitteilung.

[Wür96] K. Würde, Dissertation (Münster, 1996).

[Xu93] Y.-N. Xu und W. Y. Ching, Phys. Rev. B **48**, 4335 (1993).

[Zah94] D. R. T. Zahn, G. Kudlek, U. Rossow, A. Hoffmann, I. Broser und W. Richter, Advanced Materials for Optics and Electronics **3**, 11 (1994).

[Zak94] O. Zakharov, A. Rubio, X. Blase, M. L. Cohen und S. G. Louie, Phys. Rev. B **50**, 10 780 (1994).

[Zak95] O. Zakharov, A. Rubio und M. L. Cohen, Phys. Rev. B **51**, 4926 (1995).

[Zan89] H. J. W. Zandvliet, A. van Silfhout und M. J. Sparnaay, Phys. Rev. B **39**, 5576 (1989).

[Zhu89a] X. Zhu, S. B. Zhang, S. G. Louie und M. L. Cohen, Phys. Rev. Lett. **63**, 2112 (1989).

[Zhu89b] Z. Zhu, N. Shima und M. Tsukada, Phys. Rev. B **40**, 11 868 (1989).

[Zhu91a] X. Zhu und S. G. Louie, Phys. Rev. B **43**, 12 146 (1991).

[Zhu91b] X. Zhu und S. G. Louie, Phys. Rev. B **43**, 14 142 (1991).

An dieser Stelle möchte ich allen danken, die zum Gelingen dieser Arbeit beigetragen haben.
Ich danke vor allem Herrn Prof. Dr. J. Pollmann, der mich zu diesem Thema anregte, für sein großes Interesse an dieser Arbeit, für seine freundliche Unterstützung und seine ständige Hilfsbereitschaft.
Herrn Priv.-Doz. Dr. P. Krüger und Herrn Dr. A. Mazur gebühren großer Dank und Anerkennung für zahllose wertvolle Hinweise und Denkanstöße und für unschätzbare Diskussionen.
Ich danke Herrn K. Würde und Herrn D. Vogel für ihre Hilfsbereitschaft und für viele interessante Diskussionen.
Ferner danke ich meinen Zimmerkollegen Herrn V. Gräschus und Herrn M. Aakolk für die angenehme und kreative Zusammenarbeit.
Außerdem gilt mein Dank allen Mitarbeitern des Instituts für Theoretische Physik II für das gute Arbeitsklima und für zahlreiche wissenschaftliche (und nichtwissenschaftliche) Gespräche.
Ganz besonders danke ich meiner Freundin und Weggefährtin Magdalena Sabisch, auf deren Ermutigung, Unterstützung und Verständnis ich immer zählen konnte.

Lebenslauf

Persönliche Daten:

Name:	Michael Rohlfing
Geburtsdatum/-ort:	29. Oktober 1968 in Minden (Westf.)
Familienstand:	ledig
Eltern:	Jürgen Rohlfing Renate Rohlfing, geb. Hoppmann

Schulbildung:

1975 – 1979	Grundschule Minden-Bölhorst/Häverstädt
1979 – 1988	Ratsgymnasium Minden
14. Juni 1988	Allgemeine Hochschulreife

Studium:

10/88 – 09/93	Diplomstudiengang Physik an der Westfälischen Wilhelms-Universität Münster
13. September 1993	Diplom im Fach Physik an der WWU Münster

Promotion:

seit 10/93	Anfertigung einer Dissertation im Institut für Theoretische Physik II an der WWU Münster unter Betreuung von Prof. Dr. J. Pollmann

Tätigkeiten

04/92 – 09/93	Studentische Hilfskraft am Institut für Theoretische Physik II an der WWU Münster
seit 10/93	Wissenschaftlicher Mitarbeiter am Institut für Theoretische Physik II an der WWU Münster

. . . . . .